GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 14:01:35 ; Search time 184 Seconds
(without alignments)
3941.886 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2349554 seqs, 459636880 residues

Total number of hits satisfying chosen parameters: 4699108

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	80.4	10.2	245	5	US-09-531-113-28669	Sequence 28669, A
2	77.8	9.9	268	5	US-09-531-113-1640	Sequence 1640, Ap
c 3	76.4	9.7	247	5	US-09-531-113-3661	Sequence 3661, Ap
4	75.8	9.6	516	5	US-09-743-247A-49	Sequence 49, Appl
5	75.8	9.6	673	5	US-09-743-247A-59	Sequence 59, Appl
6	70.8	9.0	271	5	US-09-531-113-39038	Sequence 39038, A
c 7	66	8.4	528	5	US-09-531-113-29322	Sequence 29322, A
8	62.2	7.9	1081	6	US-10-223-089-333	Sequence 333, App
9	54.4	6.9	747	5	US-09-724-676-3124	Sequence 3124, Ap
10	54.4	6.9	747	5	US-09-724-676A-3124	Sequence 3124, Ap
11	54.4	6.9	771	5	US-09-724-676-3128	Sequence 3128, Ap
12	54.4	6.9	771	5	US-09-724-676A-3128	Sequence 3128, Ap
13	40	5.1	241	5	US-09-531-113-35326	Sequence 35326, A
14	38.4	4.9	8801	6	US-10-240-453-159	Sequence 159, App
15	38	4.8	6826	6	US-10-240-485-198	Sequence 198, App
c 16	37.4	4.7	598	5	US-09-531-113-29794	Sequence 29794, A
17	37.4	4.7	8201	6	US-10-240-485-64	Sequence 64, Appl
18	36.8	4.7	6078	6	US-10-240-453-195	Sequence 195, App
c 19	36.2	4.6	4673	7	US-60-423-552-63	Sequence 63, Appl
c 20	36.2	4.6	4673	7	US-60-427-579-63	Sequence 63, Appl
21	36.2	4.6	5367	6	US-10-240-485-98	Sequence 98, Appl
22	36.2	4.6	5718	5	US-09-724-676-8158	Sequence 8158, Ap
23	36.2	4.6	5718	5	US-09-724-676A-8158	Sequence 8158, Ap
24	36.2	4.6	5754	5	US-09-724-676-8120	Sequence 8120, Ap
25	36.2	4.6	5754	5	US-09-724-676A-8120	Sequence 8120, Ap
26	36.2	4.6	5815	5	US-09-724-676-8160	Sequence 8160, Ap
27	36.2	4.6	5815	5	US-09-724-676A-8160	Sequence 8160, Ap
28	36.2	4.6	5827	5	US-09-724-676-8164	Sequence 8164, Ap
29	36.2	4.6	5827	5	US-09-724-676A-8164	Sequence 8164, Ap
30	36.2	4.6	5842	5	US-09-724-676-8133	Sequence 8133, Ap
31	36.2	4.6	5842	5	US-09-724-676A-8133	Sequence 8133, Ap
32	36.2	4.6	5851	5	US-09-724-676-8122	Sequence 8122, Ap
33	36.2	4.6	5851	5	US-09-724-676A-8122	Sequence 8122, Ap
34	36.2	4.6	5863	5	US-09-724-676-8125	Sequence 8125, Ap
35	36.2	4.6	5863	5	US-09-724-676A-8125	Sequence 8125, Ap
36	36.2	4.6	5878	5	US-09-724-676-8151	Sequence 8151, Ap
37	36.2	4.6	5878	5	US-09-724-676A-8151	Sequence 8151, Ap
38	36.2	4.6	5924	5	US-09-724-676-8166	Sequence 8166, Ap
39	36.2	4.6	5924	5	US-09-724-676A-8166	Sequence 8166, Ap
40	36.2	4.6	5931	5	US-09-724-676-8145	Sequence 8145, Ap
41	36.2	4.6	5931	5	US-09-724-676A-8145	Sequence 8145, Ap
42	36.2	4.6	5939	5	US-09-724-676-8135	Sequence 8135, Ap
43	36.2	4.6	5939	5	US-09-724-676A-8135	Sequence 8135, Ap
44	36.2	4.6	5951	5	US-09-724-676-8138	Sequence 8138, Ap
45	36.2	4.6	5951	5	US-09-724-676A-8138	Sequence 8138, Ap

ALIGNMENTS

RESULT 1
US-09-531-113-28669
; Sequence 28669, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 28669
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700944276H1
US-09-531-113-28669

Query Match 10.2%; Score 80.4; DB 5; Length 245;
Best Local Similarity 67.1%; Pred. No. 5.8e-13;
Matches 114; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

```
Qy      70 ATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGAGGCATTA 129
           |||||| |  ||| |||||| ||||  ||||||   ||||| || || || ||||| ||
Db      11 ATGAGTGCATCAACCAAAGCTTGGATTGTGGCATCAAGCATAGGGGCAGTGGAGGCCTTG 70

Qy     130 AAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACATCTC 189
           ||||||  || ||||| | ||  | ||||||||     | ||||  |  | ||||||
Db      71 AAAGACCAATTAGGTGTATGCAGATGGAACTATGCTTTAAGGTCACTACAACAACATGCA 130

Qy     190 CGCAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTC 239
           |||||| | ||||||  |  |||||  |  |  || || ||  | ||
Db     131 AAGAACAACATCAGATCTTATAGTCAAGCTAGGAAATTGTCCTCTGCATC 180
```

RESULT 2
US-09-531-113-1640
; Sequence 1640, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 1640
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700959201H1
US-09-[illegible]-113-1640

Query Match 9.9%; Score 77.8; DB 5; Length 268;
Best Local Similarity 56.6%; Pred. No. 3.2e-12;
Matches 142; Conservative 1; Mismatches 108; Indels 0; Gaps 0;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          ||  |||| || ||| |||| |     ||| ||||||||||        |||||  |   |
Db     19 GAGATCACGGCGGAGGAGCTCAAGGCCTACGACGGCACCGATCCCGAGAAGCCCTTGGTC 77

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| | ||   | |  ||:||| ||    | |     || || ||  | ||
Db     78 ATGGCCATCAAGGCCCAGATCTATGATGTCTCCCACAGCAGGATGTTTTATGGACCTGGT 137

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||   |||  |  | || || || || || || |||||||| || |  ||||||  |
Db    138 GGACCTTATGCCCTATTTGCTGGCAAGGATGCTAGCAGAGCCTTAGCGAAGATGTCTTTT 197

Qy    517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTT 576
          | ||  ||||| || ||     | ||   ||||||         |||||  || |  |
Db    198 GAAGAGAAAGATCTGACTGGGGATATTTCTGGTCTCGGCCCATTCGAGATTGATGCCTTA 257

Qy    577 AATGATTGGGA 587
          | || |||||
Db    238 CAAGACTGGGA 268
```

RESULT 3
US-09-531-113-3661/c
; Sequence 3661, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 3661
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: jC-gmfl02220057c11a1
US-09-531-113-3661

Query Match 9.7%; Score 76.4; DB 5; Length 247;
Best Local Similarity 68.0%; Pred. No. 7.7e-12;
Matches 121; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

```
Qy    461 ACTCGATGTTCGCCGGAAAAGACGCGAGCA-GAGCTTTGGGTAAGATGAGTAAGAACGAA 522
          || | |||||||||||| | | |||| | || | | ||  | |  |  ||||| ||||||
Db    247 ACACCATGTTCGCCGGCACATACGCCACCAGGCGCCCTCGCCAGCATGAGCAAGAACTAC 188

Qy    523 GAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGAT 582
          || ||  | || || || || || || || ||  |  | || ||||||     ||| || ||
Db    187 GACGACATCTCCCCCTCCCTCGACGGCCTATCCTACAAGGAGATCGGCGTTCTCAACGAC 128

Qy    583 TGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCT 640
          ||||||| ||||||  |||||||||| |||||||||  ||| || ||  | |  | |
Db    127 TGGGAGAACAAATTCCAAGCTAAGTACCCTGTCGTTGCTCGTCTTCTCAATGACCTGT 70
```

RESULT 4
US-09-743-247A-49
; Sequence 49, Application US/09743247A
; GENERAL INFORMATION:
; APPLICANT: Sagami Chemical Research Center; Protegene Inc.
; TITLE OF INVENTION: Human Proteins Having Hydrophobic Domains And DNAs Encoding The
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 1997.13300
; CURRENT APPLICATION NUMBER: US/09/743,247A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: JP 10-208820
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: JP 10-224105
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: JP 10-238116
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: JP 10-254736
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: JP 10-275505
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Windows 95 (Word 98)
; SEQ ID NO 49
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-743-247A-49

Query Match 9.6%; Score 75.8; DB 5; Length 516;
Best Local Similarity 55.6%; Pred. No. 1.5e-11;
Matches 143; Conservative 1; Mismatches 113; Indels 0; Gaps 0;

```
Qy    340 TTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTC 399
          |||||||  ||| ||||   ||  ||   |||    || |||    |||| ||||||
Db    139 TTCACCGAGGAGGAGCTGGCCCGCTATGGCGGGGAGGAGGAAGATCAGCCCATCTACTTG 198

Qy    400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
          ||| | || ||    ||||| ||:|||||| |||||||    || || ||    |||| |
Db    199 GCAGTGAAGGGAGTGGTGTTTGATGTCACCTCCGGAAAGGAGTTTTATGGACGAGGAGCC 258

Qy    460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
          |||      ||  | || || ||| | |  ||||   | |  ||||||      | |
Db    259 CCCTACAATGCCTTGACGGGGAAGGACTCCACTAGAGGGGTAGCCAAGATGTCCTTGGAT 318

Qy    520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAAT 579
          | |||  |  | | |    |  |||||||| |  || ||  |  |  | ||  ||
Db    319 CCTGCAGACCTCACCCATGACACTACGGGTCTCACGGCCAAGGAACTGGAGGCCCTGGAT 378

Qy    580 GATTGGGAGACCAAATT 596
          ||       |||||| |
Db    379 GAGGTCTTCACCAAAGT 395
```

RESULT 5
US-09-743-247A-59
; Sequence 59, Application US/09743247A
; GENERAL INFORMATION:
; APPLICANT: Sagami Chemical Research Center; Protegene Inc.
; TITLE OF INVENTION: Human Proteins Having Hydrophobic Domains And DNAs Encoding The
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 1997.13300
; CURRENT APPLICATION NUMBER: US/09/743,247A
; CURRENT FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: JP 10-208820
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: JP 10-224105
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: JP 10-238116
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: JP 10-254736
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: JP 10-275505
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Windows 95 (Word 98)
; SEQ ID NO 59

; LENGTH: 673
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)...(543)
US-09-743-247A-59

Query Match 9.6%; Score 75.8; DB 5; Length 673;
Best Local Similarity 55.6%; Pred. No. 1.6e-11;
Matches 143; Conservative 1; Mismatches 113; Indels 0; Gaps 0;

Qy 340 TTCACCGCAGAGCAGCTAAGCCAATACAACCGCACCCACGAATCAAAGCCGATCTACGTC 399
Db 163 TTCACCGAGGAGGAGCTGGCCCGCTATGGCGGGGAGGAGGAAGATCAGCCCATCTACTTG 222

Qy 400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
Db 223 GCAGTGAAGGGAGTGGTGTTTGATGTCACCTCCGGAAAGGAGTTTTATGGACGAGGAGCC 282

Qy 460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
Db 283 CCCTACAATGCCTTGACGGGGAAGGACTCCACTAGAGGGGTAGCCAAGATGTCCTTGGAT 342

Qy 520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTTAAT 579
Db 343 CCTGCAGACCTCACCCATGACACTACGGGTCTCACGGCCAAGGAACTGGAGGCCCTGGAT 402

Qy 580 GATTGGGAGACCAAATT 596
Db 403 GAGGTCTTCACCAAAGT 419

RESULT 6
US-09-531-113-39038
; Sequence 39038, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 39038
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700943641H1
US-09-531-113-39038

Query Match 9.0%; Score 70.8; DB 5; Length 271;
Best Local Similarity 65.1%; Pred. No. 2.9e-10;
Matches 121; Conservative 0; Mismatches 62; Indels 3; Gaps 1;

Qy 57 TCTCTGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGC 116
Db 61 TCACAATAACAATATGAGTTCAGGAATTAGAGCATGGAGTGTGGCAACTAGCGTTGGAGT 120

Qy 117 CGTAGAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGT 176
Db 121 TGTGGAGGCCTTAAAAGACCA---AGGTATATGTAGGTGGAATCATGCTTTAAGATCAGC 177

Qy 177 TAATCAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTC 236
Db 178 TCAACACCATCTCAAAAACCATGTTGGTTCATTCTCTCAGGCTAACAAGCTTTCTTCTTC 237

Qy 237 TTCTGT 242
Db 238 TGCTAT 243

RESULT 7
US-09-531-113-29322/c
; Sequence 29322, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 29322
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: jC-gmst02400009h05d1
US-09-531-113-29322

Query Match 8.4%; Score 66; DB 5; Length 528;
Best Local Similarity 63.4%; Pred. No. 8.2e-09;
Matches 118; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

Qy 57 TCTCTGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGC 116
Db 488 TCACAATAACAAGATGAGTTCAGGAATTAGAGCATGGAGTGTGGCAGCAAGTGTTGGAAT 429

Qy 117 CGTAGAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGT 176
Db 428 TGTGGAGGCCTTGAAAGACCA---AGGCATATGCAGGTGGAATCATGCCTTAAGATCAGC 372

Qy 177 TAATCAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTC 236
Db 371 TCAACACCATCTCAAAAACCATGTTGGCTCATTCTCTCAGGCTAACAAGCTTTCTTCTTC 312

Qy 237 TTCTGT 242
Db 311 TGCTAT 306

RESULT 8
US-10-223-089-333
; Sequence 333, Application US/10223089
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235P1C9
; CURRENT APPLICATION NUMBER: US/10/223,089
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624

```
;  PRIO  FILING DATE: 2000-07-25
;  PRIO  APPLICATION NUMBER: US 60/220,664
;  PRIO  FILING DATE: 2000-07-25
;  PRIO  APPLICATION NUMBER: PCT/US00/20710
;  PRIO  FILING DATE: 2000-07-28
;  PRIO  APPLICATION NUMBER: US 60/222,695
;  PRIO  FILING DATE: 2000-08-02
;  PRIO  APPLICATION NUMBER: US 09/643,657
;  PRIO  FILING DATE: 2000-08-17
;  PRIO  APPLICATION NUMBER: PCT/US00/23522
;  PRIO  FILING DATE: 2000-08-23
;  PRIO  APPLICATION NUMBER: PCT/US00/23328
;  PRIO  FILING DATE: 2000-08-24
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 333
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-223-089-333

  Query Match              7.9%;  Score 62.2;  DB 6;  Length 1081;
  Best Local Similarity   54.6%;  Pred. No. 1.2e-07;
  Matches  171;  Conservative    1;  Mismatches  129;  Indels   12;  Gaps    2;

Qy    320 CTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACG 379
          ||    | ||| |||  ||| |||| |   ||||||||  |||| ||| ||||| ||  |
Db    290 CTCGCATGAAGAAGCGGGACTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCA 349

Qy    380 AATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAAT 439
          |  |    ||  |||    ||||  |||| ||    || |||||:|| ||||  .| |
Db    350 ACCC---GCGCATCCTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCA 406

Qy    440 CCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTT 499
          |||||||||  | |  ||    ||    || || || || |  || ||   |'|||
Db    407 AGTTCTACGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGAC 466

Qy    500 TGGGTAAGATGAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTC 550
          |||  |  |  |   |  ||||         |||||  | |  |  |||   || |
Db    467 TGGCCACATTTTGCCTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATT 526

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          | | ||    | | ||  | | | ||   || |||||  |   | |||  |||   | |||
Db    527 TGAATGCAGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATG 586

Qy    611 CTGTCGTTGGCCG 623
          |   || ||| |
Db    587 ATTATGTAGGCAG 599


RESULT 9
US-09-724-676-3124
; Sequence 3124, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3124
;   LENGTH: 747
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-724-676-3124

  Query Match              6.9%;  Score 54.4;  DB 5;  Length 747;
  Best Local Similarity   54.1%;  Pred. No. 1.6e-05;
  Matches  159;  Conservative    1;  Mismatches  122;  Indels   12;  Gaps    2;

Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |   ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    304 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 360

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   |||||||||:|| ||||  ||        ||||||||  |||
Db    361 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 420

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | | |   |  |  |
Db    421 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 480

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||          |||| ||    |  |||   |  |||||||       ||
Db    481 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 540

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| |||||| |  | ||       ||||||| |  ||| |||
Db    541 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 594


RESULT 10
US-09-724-676A-3124
; Sequence 3124, Application US/09724676A
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676A
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3124
;   LENGTH: 747
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-724-676A-3124

  Query Match              6.9%;  Score 54.4;  DB 5;  Length 747;
  Best Local Similarity   54.1%;  Pred. No. 1.6e-05;
  Matches  159;  Conservative    1;  Mismatches  122;  Indels   12;  Gaps    2;

Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |   ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    304 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 360

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   |||||||||:|| ||||  ||        ||||||||  |||
Db    361 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 420

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | | |   |  |  |
Db    421 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 480

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||          |||| ||    |  |||   |  |||||||       ||
Db    481 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 540

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| |||||| |  | ||       ||||||| |  ||| |||
Db    541 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 594


RESULT 11
US-09-724-676-3128
; Sequence 3128, Application US/09724676
; GENERAL INFORMATION:
;  APPLICANT: Compugen LTD
;  TITLE OF INVENTION: Variants of alternative splicing
;  FILE REFERENCE: 129181.4 Compugen
;  CURRENT APPLICATION NUMBER: US/09/724,676
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 97222
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3128
```

```
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (729)..(729)
;   OTHER INFORMATION: n is a,c,g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (756)..(756)
;   OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-3128
```

Query Match 6.9%; Score 54.4; DB 5; Length 771;
Best Local Similarity 54.1%; Pred. No. 1.7e-05;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |   |||| | |  | | ||||| | | ||  |   ||  ||   |
Db    304 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 360

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||       |||||||  |||
Db    361 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 420

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | |  |   |  |   |
Db    421 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 480

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||         |||| ||    |  |||   | |||||||      ||
Db    481 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 540

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| |||||| |  | ||       |||||| |  ||| |||
Db    541 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 594
```

RESULT 12
US-09-724-676A-3128
; Sequence 3128, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3128
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (729)..(729)
; OTHER INFORMATION: n is a,c,g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (756)..(756)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-3128

Query Match 6.9%; Score 54.4; DB 5; Length 771;
Best Local Similarity 54.1%; Pred. No. 1.7e-05;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |   |||| | |  | | ||||| | | ||  |   ||  ||   |
Db    304 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 360

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||       |||||||  |||
```

```
Db    361 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 420

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | |  |   |  |   |
Db    421 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 480

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||         |||| ||    |  |||   | |||||||      ||
Db    481 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 540

Qy    568 AATACTCTTAATGATTGGGAGACCAAA:TTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| |||||| |  | ||       |||||| |  ||| |||
Db    541 GAGACTCTGAGTGACTGGGAGTCTCA:TTCACTTTCAAGTATCATCACGTGGGC 594
```

RESULT 13
US-09-531-113-35326
; Sequence 35326, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 35326
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700890003H1
US-09-531-113-35326

Query Match 5.1%; Score 40; DB 5; Length 241;
Best Local Similarity 52.4%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy     97 GTGGCAGTGAGCATCGGAGCCGTAGAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGG 156
          || |||   |   | ||||  || |||| || || || |||     |       |
Db     52 GTTGCAACCACTGTTGGAGTGGTGGAGGCGTTGAAGGATCAAGGGTATTGCAAAATGAAC 111

Qy    157 AACTACATACTCCGGTCGGTTAATCAACATCTCCGGAACAACGTTAGATCTGTTTCTCAA 216
          | |   ||  |   || |||   |||||||  |  |||| || |  | || ||  || |
Db    112 AGCACAATGATGAAGTTGGTGGCTCAACATGCCAAGAACCACATAGGGTCAGTAACTGAG 171

Qy    217 GGGAAAAGGTTCTCTTCGTCTTCTGTCTCCGCAGCCGTTACCTCCTCT 264
          |  || | | |  ||||  |||||   ||  |  |   | | || |||
Db    172 GCCAAGAAGCTGGCTTCTCCTTCTCCTTCTCCTTCTACTTCTTCTTCT 219
```

RESULT 14
US-10-240-453-159
; Sequence 159, Application US/10240453
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA TRanscription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019178.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 350
SEQ ID NO 159
LENGTH: 8801
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
10-240-453-159

uery Match 4.9%; Score 38.4; DB 6; Length 8801;
est Local Similarity 58.9%; Pred. No. 1.2;
atches 66; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

660 TGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTG 719
2252 TGTTATGTAAAAAAGGTGTTTGTGTGTTTTTGGTTTGGGTTTTTTTTTTTTTTAAGGATTT 2311

720 GATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATA 771
2312 GTTTTTTTTTTTTTTAAAAAAATTATGTAATATAAGTAGTTAATCGTTAAAAA 2363

SULT 15
10-240-485-198
Sequence 198, Application US/10240485
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with
TITLE OF INVENTION: Metastasis
FILE REFERENCE: 5013.1007
CURRENT APPLICATION NUMBER: US/10/240,485
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/EP01/03970
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 202
SEQ ID NO 198
LENGTH: 6826
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
10-240-485-198

Query Match 4.8%; Score 38; DB 6; Length 6826;
Best Local Similarity 57.6%; Pred. No. 1.4;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

624 TGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAG 683
6161 TTTTTTATATTTTATTTCGTTTTTTTTATTTTATTTTTTTAAGTTATTTTTGTTTTTTTT 6220

684 GATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAAT 741
6221 TTTTTTTTTTTTTTTTTTTTTTTGAGTTAGTGTTTTGTTGTGTTATTTAGGTTAGAGT 6278

arch completed: January 8, 2003, 14:17:58
time : 203 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 14:03:45 ; Search time 3454 Seconds
(without alignments)
5743.319 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		788.6	99.9	789	19	US-09-513-996A-70554	Sequence 70554, A
2		788.6	99.9	789	22	US-09-565-309A-56330	Sequence 56330, A
3		788.6	99.9	789	25	US-09-649-866A-1	Sequence 1, Appli
4		788.6	99.9	792	22	US-09-565-309A-48301	Sequence 48301, A
5		614.2	77.8	631	22	US-09-565-309A-1451	Sequence 1451, Ap
6		489.2	62.0	522	19	US-09-513-996A-12108	Sequence 12108, A
7		489.2	62.0	522	24	US-09-620-394B-328	Sequence 328, App
c 8		489.2	62.0	103495	20	US-09-534-859-278	Sequence 278, App
c 9		489.2	62.0	103495	31	US-09-803-736-278	Sequence 278, App
10		478.6	60.7	656	25	US-09-654-617-126256	Sequence 126256,
11		478.6	60.7	656	27	US-09-684-016-126256	Sequence 126256,
12		477.2	60.5	521	22	US-09-565-309A-8446	Sequence 8446, Ap
13		477.2	60.5	521	22	US-09-565-309A-43324	Sequence 43324, A
14		470.6	59.6	498	30	US-09-770-961-478	Sequence 478, App
15		455.6	57.7	456	22	US-09-565-309A-1449	Sequence 1449, Ap
c 16		409	51.8	436	25	US-09-649-165A-4119	Sequence 4119, Ap
17		402.2	51.0	474	19	US-09-513-996A-1123	Sequence 1123, Ap
18		402.2	51.0	474	22	US-09-565-309A-5609	Sequence 5609, Ap
19		402.2	51.0	474	22	US-09-565-309A-62352	Sequence 62352, A
20		402.2	51.0	474	22	US-09-595-335A-373	Sequence 373, App
21		402.2	51.0	482	22	US-09-565-309A-5608	Sequence 5608, Ap

22	402.2	51.0	489	22	US-09-565-309A-5610	Sequence 5610, Ap
23	401.8	50.9	474	22	US-09-565-309A-53681	Sequence 53681, A
24	390.2	49.5	448	22	US-09-565-309A-5607	Sequence 5607, Ap
25	366.6	46.5	367	22	US-09-565-309A-1450	Sequence 1450, Ap
26	299.6	38.0	300	22	US-09-565-309A-67871	Sequence 67871, A
27	282.8	35.8	417	22	US-09-565-309A-1375	Sequence 1375, Ap
28	282.8	35.8	477	19	US-09-513-996A-8419	Sequence 8419, Ap
29	282.8	35.8	477	22	US-09-565-309A-1377	Sequence 1377, Ap
30	282.8	35.8	477	22	US-09-565-309A-48124	Sequence 48124, A
31	282.8	35.8	477	22	US-09-565-309A-60595	Sequence 60595, A
32	282.8	35.8	493	22	US-09-565-309A-1376	Sequence 1376, Ap
33	282.8	35.8	83922	31	US-09-803-736-1386	Sequence 1386, Ap
34	282.8	35.8	99856	20	US-09-534-859-562	Sequence 562, App
35	282.8	35.8	99856	31	US-09-803-736-562	Sequence 562, App
36	282.4	35.8	493	22	US-09-565-309A-1378	Sequence 1378, Ap
37	276.8	35.1	413	19	US-09-513-996A-15206	Sequence 15206, A
38	276.8	35.1	413	22	US-09-565-309A-46983	Sequence 46983, A
39	276.8	35.1	413	24	US-09-620-393B-3290	Sequence 3290, Ap
40	276.8	35.1	417	22	US-09-565-309A-42300	Sequence 42300, A
41	272.8	34.6	420	25	US-09-649-165A-7293	Sequence 7293, Ap
42	213.8	27.1	270	22	US-09-565-309A-64086	Sequence 64086, A
43	195.6	24.8	485	25	US-09-654-617-9393	Sequence 9393, Ap
44	195.6	24.8	485	27	US-09-684-016-9393	Sequence 9393, Ap
45	194.6	24.7	419	17	US-09-304-517A-257238	Sequence 257238,

ALIGNMENTS

```
RESULT 1
US-09-513-996A-70554
 Sequence 70554, Application US/09513996A
 GENERAL INFORMATION:
  APPLICANT: N. ALEXANDROV et al.
  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
  TITLE OF INVENTION:  ENCODED THEREBY
  FILE REFERENCE: 2750-709P
  CURRENT APPLICATION NUMBER: US/09/513,996A
  CURRENT FILING DATE:  2000-02-25
  NUMBER OF SEQ ID NOS: 81028
 SEQ ID NO 70554
   LENGTH: 789
   TYPE: DNA
   ORGANISM: Arabidopsis thaliana
   FEATURE:
   NAME/KEY: UNSURE
   LOCATION: 1..789
   OTHER INFORMATION: any n or Xaa = unknown
   FEATURE:
   OTHER INFORMATION: Location 1..789 / Ceres Seq. ID 2218850
US-09-513-996A-70554

 Query Match             99.9%;  Score 788.6;  DB 19;  Length 789;
 Best Local Similarity   100.0%;  Pred. No. 2.3e-199;
 Matches  789;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300

Qy     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480

Qy     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540

Qy     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600

Qy     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660

Qy     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720

Qy     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780

Qy     781 TTTCGTGTT 789
           |||||||||
Db     781 TTTCGTGTT 789


RESULT 2
US-09-565-309A-56330
; Sequence 56330, Application US/09565309A
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai
;  APPLICANT:  BROVER, Vyacheslav
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTI
;  TITLE OF INVENTION:  THEREBY
;  FILE REFERENCE: 2750-0853P
;  CURRENT APPLICATION NUMBER: US/09/565,309A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 56330
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(789)
;   OTHER INFORMATION: any n = a, g, c, t, unknown, or other
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(789)
;   OTHER INFORMATION: 1916  : OVERLAP (Clone Number : OVERLAP)
US-09-565-309A-56330

  Query Match             99.9%;  Score 788.6;  DB 22;  Length 789;
  Best Local Similarity   99.9%;  Pred. No. 2.3e-199;
  Matches  788;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
```

```
Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300

Qy     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480
           ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480

Qy     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540

Qy     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600

Qy     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660

Qy     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720

Qy     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780

Qy     781 TTTCGTGTT 789
           |||||||||
Db     781 TTTCGTGTT 789
```

RESULT 3
US-09-649-866A-1
; Sequence 1, Application US/09649866A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1097P
; CURRENT APPLICATION NUMBER: US/09/649,866A
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 3537
; SEQ ID NO 1
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..789
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc feature
; LOCATION: 1..789
; OTHER INFORMATION: Ceres Seq. ID 1497851
US-09-649-866A-1

Query Match 99.9%; Score 788.6; DB 25; Length 789;
Best Local Similarity 100.0%; Pred. No. 2.3e-199;
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60

Qy      61 TGAAGAAAGA AGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGA AGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy     181 CAACATCTCC GAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACATCTCC GAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy     241 GTCTCCGCAG CGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTCTCCGCAG CGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300

Qy     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy     421 GAYGTCACCA CGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAYGTCACCA CGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480

Qy     481 AAAGACGCGA CAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAAGACGCGA CAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540

Qy     541 CTTGAAGGTC CACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTTGAAGGTC CACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600

Qy     601 GCTAAGTATC TGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTAAGTATC TGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660

Qy     661 GTTATGTAAC ATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTATGTAAC ATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720

Qy     721 ATCTGATCGT TTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ATCTGATCGT TTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780

Qy     781 TTTCGTGTT 789
           |||||||||
Db     781 TTTCGTGTT 789
```

RESULT 4
US-09-565-309A-48301
; Sequence 48301, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES I
; TITLE OF INVENTION: THEREBY

FILE REFERENCE: 2750-0853P
CURRENT APPLICATION NUMBER: US/09/565,309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 68449
SEQ ID NO 48301
LENGTH: 792
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(792)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
NAME/KEY: misc_feature
LOCATION: (1)..(792)
OTHER INFORMATION: 1916 : 5TAG CONSENSUS (Clone Number:5tag_consensus)
US-09-565-309A-48301

Query Match 99.9%; Score 788.6; DB 22; Length 792;
Best Local Similarity 99.9%; Pred. No. 2.3e-199;
Matches 788; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60

Qy     61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy    121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy    181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy    241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCGCTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCGCTT 300

Qy    301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy    361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy    421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480
          ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480

Qy    481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540

Qy    541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600

Qy    601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660

Qy    661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720

Qy    721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
```

```
Qy    781 TTTCGTGTT 789
          |||||||||
Db    781 TTTCGTGTT 789
```

RESULT 5
US-09-565-309A-1451
; Sequence 1451, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 1451
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(631)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(631)
; OTHER INFORMATION: 1916:146992 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-1451

Query Match 77.8%; Score 614.2; DB 22; Length 631;
Best Local Similarity 99.7%; Pred. No. 7.1e-153;
Matches 613; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy    175 GTTAATCAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCG 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTTAATCAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCG 60

Qy    235 TCTTCTGTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTT 294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTTCTGTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTT 120

Qy    295 TCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAG 354
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCCCTTGAGAWACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAG 180

Qy    355 CTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGT 414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGT 240

Qy    415 GTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTC 474
          ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTC 300

Qy    475 GCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCT 534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCT 360

Qy    535 CCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAA 594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAA 420

Qy    595 TTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTG 654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTG 480

Qy    655 CACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGT 714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGT 540

Qy    715 GTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAAT 774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAAT 600
```

```
Qy     775 CGGGGATTTCGTGTT 789
           ||||||||||||||| 
Db     601 CGGGGATTTCGTGTT 615


RESULT 6
US-09-513-996A-12108
; Sequence 12108, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 12108
;   LENGTH: 522
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..522
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..522 / Ceres Seq. ID 1376054
US-09-513-996A-12108

  Query Match             62.0%;  Score 489.2;  DB 19;  Length 522;
  Best Local Similarity   99.0%;  Pred. No. 1.5e-119;
  Matches  491;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 73

Qy     354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy     414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
           |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 193

Qy     474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 253

Qy     534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 313

Qy     594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     314 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 373

Qy     654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     374 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 433

Qy     714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     434 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 493

Qy     774 TCGGGGATTTCGTGTT 789
           ||||||||||||||| 
Db     494 TCGGGGATTTCGTGTT 509


RESULT 7
US-09-620-394B-328
; Sequence 328, Application US/09620394B
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai
;  APPLICANT:  BROVER, Vyacheslav
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 2750-1067P
;  CURRENT APPLICATION NUMBER: US/09/620,394B
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 328
;   LENGTH: 522
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1..522
;   OTHER INFORMATION: any n = a, g, c, t, unknown, or other
;   NAME/KEY: misc_feature
;   LOCATION: 1..522
;   OTHER INFORMATION: Ceres Seq. ID 1376054
US-09-620-394B-328

  Query Match             62.0%;  Score 489.2;  DB 24;  Length 522;
  Best Local Similarity   99.0%;  Pred. No. 1.5e-119;
  Matches  491;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 73

Qy     354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy     414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
           |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 193

Qy     474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 253

Qy     534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 313

Qy     594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     314 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 373

Qy     654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     374 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 433

Qy     714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     434 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 493

Qy     774 TCGGGGATTTCGTGTT 789
           ||||||||||||||| 
Db     494 TCGGGGATTTCGTGTT 509


RESULT 8
US-09-534-859-278
; Sequence 278, Application US/09534859
; GENERAL INFORMATION:
;  APPLICANT:       Bush, David F.
;  APPLICANT:  Last, Robert L.
;  APPLICANT:  Lehin, Irena M.
;  APPLICANT:  Norris, Susan R.
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Rounsley, Steven D.
```

```
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION:       PLANT POLYMORPHIC MARKERS AND USES THEREOF
;  FILE REFERENCE:       38-10(15493)B
;  CURRENT APPLICATION NUMBER: US/09/534,859
;  CURRENT FILING DATE:  2000-03-29
;  NUMBER OF SEQ ID NOS:       1127
; SEQ ID NO 278
;   LENGTH: 103495
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-534-859-278

  Query Match             62.0%;  Score 489.2;  DB 20;  Length 103495;
  Best Local Similarity   99.0%;  Pred. No. 1.5e-118;
  Matches  491;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57585 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 57526

Qy     354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57525 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 57466

Qy     414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
           |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57465 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 57406

Qy     474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57405 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 57346

Qy     534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57345 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 57286

Qy     594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57285 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 57226

Qy     654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57225 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 57166

Qy     714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57165 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 5 106

Qy     774 TCGGGGATTTCGTGTT 789
           ||||||||||||||||
Db   57105 TCGGGGATTTCGTGTT 57090


RESULT 9
US-09-803-736-278/c
; Sequence 278, Application US/09803736
; GENERAL INFORMATION:
;  APPLICANT: Bush, David F.
;  APPLICANT:  Levin, Irena M.
;  APPLICANT:  Norris, Susan R.
;  APPLICANT:  Rounsley, Steven D.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
;  FILE REFERENCE: 38-10(15493)D
;  CURRENT APPLICATION NUMBER: US/09/803,736
;  CURRENT FILING DATE:  2001-03-12
;  PRIOR APPLICATION NUMBER: US 09/534,859
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: identified by Attorney Docket number 04983.0206CPUS01 38-10
;  PRIOR FILING DATE: 2000-10-20
;  NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 278
;   LENGTH: 103495
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-803-736-278

  Query Match             62.0%;  Score 489.2;  DB 31;  Length 103495;
  Best Local Similarity   99.0%;  Pred. No. 1.5e-118;
  Matches  491;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57585 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 57526

Qy     354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57525 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 57466

Qy     414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
           |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57465 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 57406

Qy     474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57405 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 57346

Qy     534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57345 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 57286

Qy     594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57285 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 57226

Qy     654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57225 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 57166

Qy     714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   57165 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 57106

Qy     774 TCGGGGATTTCGTGTT 789
           ||||||||||||||||
Db   57105 TCGGGGATTTCGTGTT 57090


RESULT 10
US-09-654-617-126256
; Sequence 126256, Application US/09654617
; GENERAL INFORMATION:
;  APPLICANT:       Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  TITLE OF INVENTION:       Annotated Plant Genes
;  FILE REFERENCE:       38-21(15097)D
;  CURRENT APPLICATION NUMBER: US/09/654,617
;  CURRENT FILING DATE:  2000-09-05
;  NUMBER OF SEQ ID NOS:       463173
; SEQ ID NO 126256
;   LENGTH: 656
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   OTHER INFORMATION: unsure at all n locations .
US-09-654-617-126256

  Query Match             60.7%;  Score 478.6;  DB 25;  Length 656;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-116;
  Matches  478;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     311 TGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACG 370
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40 TGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACG 99

Qy     371 GCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCA 430
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
```

```
Db    100 GCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGATGTCACCA 159

Qy    431 CCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGA 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 CCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGA 219

Qy    491 GCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTC 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 GCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTC 279

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 339

Qy    611 CTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAAC 670
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    340 CTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAAC 399

Qy    671 TATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGT 730
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400 TATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGT 459

Qy    731 TTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 TTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 518
```

RESULT 11
US-09-684-016-126256
; Sequence 126256, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 126256
; LENGTH: 656
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(656)
; OTHER INFORMATION: unsure at all n locations
US-09-684-016-126256

Query Match 60.7%; Score 478.6; DB 27; Length 656;
Best Local Similarity 99.8%; Pred. No. 1.1e-116;
Matches 478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy    311 TGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACG 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 TGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACG 99

Qy    371 GCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCA 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db    100 GCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGATGTCACCA 159

Qy    431 CCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGA 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 CCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGA 219

Qy    491 GCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTC 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 GCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTC 279

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 339
```

```
Qy    611 CTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAAC 670
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    340 CTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAAC 399

Qy    671 TATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGT 730
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400 TATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGT 459

Qy    731 TTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 789
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 TTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 518
```

RESULT 12
US-09-565-309A-8446
; Sequence 8446, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 8446
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(521)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(521)
; OTHER INFORMATION: 10261:44974 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-8446

Query Match 60.5%; Score 477.2; DB 22; Length 521;
Best Local Similarity 98.8%; Pred. No. 2.4e-116;
Matches 490; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```
Qy    294 TTCCCTTGAAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
          ||| | |..||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 73

Qy    354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy    414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
          |: ||||||:|||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    134 TTGGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGA-GCGATTACTCGATGTT 192

Qy    474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
          |.||||||.||||||||||||||||||||||||||||||||||||||||||||||||||:
Db    193 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 252

Qy    534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
          ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 312

Qy    594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 372

Qy    654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 432

Qy    714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
          |||||||||:|||||||||||||||||||||||||||||||||||||||||||||..|||
```

```
Db    433 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 492

Qy    774 TCGGGGATTTCGTGTT 789
          ||||||||||||||||
Db    493 TCGGGGATTTCGTGTT 508
```

RESULT 13
US-09-565-309A-43324
; Sequence 43324, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 43324
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(521)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(521)
; OTHER INFORMATION: 10261 : 5TAG CONSENSUS (Clone Number:5tag_consensus)
US-09-565-309A-43324

Query Match 60.5%; Score 477.2; DB 22; Length 521;
Best Local Similarity 98.8%; Pred. No. 2.4e-116;
Matches 490; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```
Qy    294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
          |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 73

Qy    354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy    414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
          |||||||||:|||||||||||||||||||||||||||||| ||||||||||||||||||
Db    134 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGA-GCGATTACTCGATGTT 192

Qy    474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 252

Qy    534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 312

Qy    594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 372

Qy    654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 432

Qy    714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 492

Qy    774 TCGGGGATTTCGTGTT 789
          ||||||||||||||||
Db    493 TCGGGGATTTCGTGTT 508
```

RESULT 14
US-09-770-961-478
; Sequence 478, Application US/09770961
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2026 (PARA-015PRV)
; CURRENT APPLICATION NUMBER: US/09/770,961
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,466
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-961-478

Query Match 59.6%; Score 470.6; DB 30; Length 498;
Best Local Similarity 99.8%; Pred. No. 1.3e-114;
Matches 470; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy    319 GCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGAC 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 GCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGAC 69

Qy    379 GAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAA 438
          ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db     70 GAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGATGTCACCACCGGAAAA 129

Qy    439 TCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCT 498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 TCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCT 189

Qy    499 TTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAG 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 TTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAG 249

Qy    559 AAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTT 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 AAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTT 309

Qy    619 GGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGT 678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 GGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGT 369

Qy    679 GTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATAC 738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 GTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATAC 429

Qy    739 AATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 789
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 AATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 480
```

```
RESULT 15
US-09-565-309A-1449
; Sequence 1449, Application US/09565309A
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai
;  APPLICANT:  BROVER, Vyacheslav
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  THEREBY
;  FILE REFERENCE: 2750-0853P
;  CURRENT APPLICATION NUMBER: US/09/565,309A
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 1449
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(456)
;   OTHER INFORMATION: any n = a, g, c, t, unknown, or other
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(456)
;   OTHER INFORMATION: 1916:36727 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-1449

  Query Match            57.7%;  Score 455.6;  DB 22;  Length 456;
  Best Local Similarity  99.8%;  Pred. No. 1.3e-110;
  Matches  455;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300

Qy     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
           ||:||||||||||||||||||||||||||||||||
Db     421 GATGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456


Search completed: January  8, 2003, 15:19:42
Job time : 3650 secs
```

```
Db    451 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 504


RESULT 9
US-09-880-107-3858
; Sequence 3858, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3858
;   LENGTH: 1890
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 Y12711.
US-09-880-107-3858

  Query Match             6.9%;  Score 54.4;  DB 10;  Length 1890;
  Best Local Similarity   54.1%;  Pred. No. 9.4e-06;
  Matches  159;  Conservative    1;  Mismatches  122;  Indels   12;  Gaps    2;

Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |   ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    292 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 348

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||       ||||||||  |||
Db    349 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 408

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | |  |   |  | |
Db    409 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 468

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||          |||| ||    |   |||   |  |||||||      ||
Db    469 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 528

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| |||||| |  | ||       |||||| |  ||| |||
Db    529 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 582


RESULT 10
US-09-783-590-11410
; Sequence 11410, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11410
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (38)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (48)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (134)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (211)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (223)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (301)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (324)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (327)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (351)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (355)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (377)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (384)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (398)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11410

  Query Match             5.7%;  Score 44.8;  DB 10;  Length 415;
  Best Local Similarity   57.8%;  Pred. No. 0.0025;
  Matches   93;  Conservative    1;  Mismatches   66;  Indels    1;  Gaps    1;

Qy    336 GGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTA 395
          | | ;|||| | |   |||  |    ||| |||||  | | | | |   | | ||
Db     97 GAACTTCACCCCCGCCGAGCTGCGGGCTTACGACGGCNTCCAGGGACCCGCGGC-ATACT 155

Qy    396 CGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGG 455
          | | || ||||| |||   ||||||||:|| ||||  ||       ||||||||  |||
Db    156 CATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGNCCGA 215

Qy    456 AGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAG 496
          ||   | |   | | || || |||| ||  ||   ||| |
Db    216 GGGGCGTNATGGGGTCTTTGCTGGAAGAGGTGCATCCAGGG 256


RESULT 11
US-09-960-352-11743
; Sequence 11743, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 13:31:40 ; Search time 49 Seconds
(without alignments)
7086.562 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 389086 seqs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	89.6	11.4	751	10	US-09-770-149-1	Sequence 1, Appli
2	76.6	9.7	254	10	US-09-923-876-382	Sequence 382, App
3	64.8	8.2	430	10	US-09-924-035A-4	Sequence 4, Appli
4	62.2	7.9	672	9	US-10-164-871-3	Sequence 3, Appli
5	62.2	7.9	1936	9	US-10-098-841-217	Sequence 217, App
6	62.2	7.9	2776	9	US-09-984-245-78	Sequence 78, Appl
7	54.8	6.9	253	10	US-09-923-876-197	Sequence 197, App
8	54.4	6.9	588	9	US-10-164-871-1	Sequence 1, Appli
9	54.4	6.9	1890	10	US-09-880-107-3858	Sequence 3858, Ap
10	44.8	5.7	415	10	US-09-783-590-11410	Sequence 11410, A
11	44.8	5.7	435	10	US-09-960-352-11743	Sequence 11743, A
c 12	40.2	5.1	1053	10	US-09-898-416-5	Sequence 5, Appli
13	38.8	4.9	252	10	US-09-878-574-6992	Sequence 6992, Ap
14	38.6	4.9	755	10	US-09-910-943-363	Sequence 363, App
15	36.6	4.6	393	10	US-09-960-352-4582	Sequence 4582, Ap
c 16	36.2	4.6	360	9	US-10-040-739-1492	Sequence 1492, Ap
c 17	35.6	4.5	1496	10	US-09-898-416-4	Sequence 4, Appli
18	35.6	4.5	32176	10	US-09-764-869-1953	Sequence 1953, Ap
19	35.6	4.5	32178	10	US-09-764-869-1954	Sequence 1954, Ap
20	35.4	4.5	19553	10	US-09-764-847-1425	Sequence 1425, A
c 21	35	4.4	383	9	US-10-025-380-67	Sequence 67, Appl
c 22	35	4.4	383	10	US-09-922-217-67	Sequence 67, App
c 23	35	4.4	383	10	US-09-833-263-67	Sequence 67, App
24	34.8	4.4	370	10	US-09-960-352-11096	Sequence 11096,
c 25	34.6	4.4	7200	9	US-09-853-450-48	Sequence 48, Appl
c 26	34.4	4.4	1633	9	US-10-028-072-217	Sequence 217, App
27	33.8	4.3	32134	10	US-09-764-847-1057	Sequence 1057, A
28	33.8	4.3	32134	10	US-09-764-877-3535	Sequence 3535, A
29	33.6	4.3	6830	10	US-09-927-668-1	Sequence 1, Appl
c 30	33.4	4.2	312	10	US-09-960-352-8414	Sequence 8414, A
31	33.4	4.2	449	10	US-09-770-444-720	Sequence 720, Ap
32	33.4	4.2	600	12	US-10-044-090-603	Sequence 603, Ap
33	33.4	4.2	2068	12	US-10-044-090-519	Sequence 519, Ap
34	33.4	4.2	2484	12	US-10-044-090-692	Sequence 692, Ap
35	33.4	4.2	3276	9	US-09-764-868-266	Sequence 266, App
36	33.2	4.2	291	10	US-09-960-352-1243	Sequence 1243, A
37	33.2	4.2	461	10	US-09-864-761-5800	Sequence 5800, A
38	33	4.2	337	10	US-09-960-352-6976	Sequence 6976, A
c 39	33	4.2	429	10	US-09-983-965-2539	Sequence 2539, A
40	32.6	4.1	419	10	US-09-960-352-11234	Sequence 11234,
c 41	32.6	4.1	458	10	US-09-954-456-1092	Sequence 1092, A
42	32.6	4.1	879	9	US-09-938-842A-3466	Sequence 3466, Ap
43	32.6	4.1	2000	10	US-09-887-576-99	Sequence 99, App
44	32.6	4.1	2000	10	US-09-887-576-147	Sequence 147, Ap
45	32.6	4.1	2786	12	US-10-052-586-307	Sequence 307, Ap

ALIGNMENTS

RESULT 1
US-09-770-149-1
; Sequence 1, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2024 (PARA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-149-1

Query Match 11.4%; Score 89.6; DB 10; Length 751;
Best Local Similarity 56.6%; Pred. No. 4.5e-16;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;

```
     337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGC GATCTAC 396
         ||  |||| |  ||| |||| |  || ||| | ||| | ||       |||| |
     186 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAG CCTTCTT 245

     397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTAC GCTCCGGA 456
         | || ||||||   |   | |  ||:|| ||     | |     ||||||   | |||
     246 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTAC ACCAGGA 305

     457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAG TGAGTAAG 516
         ||    ||  |  |||| || || |||||||| ||| ||||| | |  |||  |
     306 GGACCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAG TGTCATTT 365

     517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC TACTCTT 576
         | ||  ||||  || ||     | |    |||||    |      ||| |   | |||||
     366 GAGGAGAAAGACTTGACTTGGGATGTCTCTGGTCTTGGTCCCTTTGAGCTA CTGCTCTT 425

     577 AATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT 626
         | |||||||||  ||| ||       ||||| ||   |||||   |||
     426 CAAGATTGGGAGTACAAGTTCATGAGCAAGTATGCTAAGGTTGGTACTGT 475


SULT 2
-09-923-876-382
Sequence 382, Application US/09923876
Patent No. US20020013958A1
GENERAL INFORMATION:
 APPLICANT: Lalgudi, Raghunath V.
 APPLICANT:  Kamigaki, Laura Y. (Ito)
 APPLICANT:  Sherman, Bradley K.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
 FILE REFERENCE: PL-0012-1 CON
 CURRENT APPLICATION NUMBER: US/09/923,876
 CURRENT FILING DATE:  2001-08-06
 PRIOR APPLICATION NUMBER: 09/298,329
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/085,331
 PRIOR FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 6332
 SOFTWARE: PERL Program
SEQ ID NO 382
  LENGTH: 254
  TYPE: DNA
  ORGANISM: Zea mays
  FEATURE:
  NAME/KEY: misc_feature
  OTHER INFORMATION: Incyte ID No. US20020013958A1 70015684381
-09-923-876-382

Query Match             9.7%;  Score 76.6;  DB 10;  Length 254;
Best Local Similarity   61.3%;  Pred. No. 1.4e-12;
Matches  155;  Conservative    1;  Mismatches   95;  Indels    2;  Gaps    2;

     396 CGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAATCCTTCTAC GCTCCGG 455
         ||||  |  ||   |||   || | |||:|||||| ||||      ||||||  || ||||
       1 CGTCTCCGTCCGGGGCAAGGTCTACGACGTCACCTCCGGCCGCGGCTTCTA GGCCCGG 60

     456 AGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAA 515
         ||||  |||  |  | ||||| ||    || || ||| | ||  | || ||  ||   ||
      61 CGGCGCCTACGCCGTCTTCGCGGGCCGCGAGGC-AGCCGCGCCCTCGGCAA TGTCCAA 119

     516 GAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC AATACTCT 575
         | |||| |  || || || |     ||    || ||||| || || ||| |       ||
     120 GGACGAGGCCGACGTCTC-CGGGGACTCTCCGGGCTCACCGACAAGGAGCT GGCGTCCT 178

     576 TAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT GTCTCTTA 635
         || ||||||||||| ||  | || ||||| || ||||| | |||  | | | |  |
     179 CGCCGACTGGGAGACCAAGTTCCAGGCCAAGTACCCCGTCGTCGCCCGACT GCCGCCGA 238

     636 GGTCTCTCTTCTG 648
         | ||    || |
     239 CGCCTGAACTCGG 251
```

```
RESULT 3
US-09-924-035A-4
; Sequence 4, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
;  APPLICANT: Grlach, Jrn
;  TITLE OF INVENTION: Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE: 2011US
;  CURRENT APPLICATION NUMBER: US/09/924,035A
;  CURRENT FILING DATE:  2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/148,784
;  PRIOR FILING DATE: 1999-08-13
;  NUMBER OF SEQ ID NOS: 900
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-924-035A-4

  Query Match             8.2%;  Score 64.8;  DB 10;  Length 430;
  Best Local Similarity   57.6%;  Pred. No. 4.5e-09;
  Matches  114;  Conservative    1;  Mismatches   83;  Indels    0;  Gaps    0;

Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          ||  |||| |  ||| |||| |  || ||| | ||| | ||       ||||||  |
Db    222 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAGCCCCTTCTT 281

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||||   |   | |  ||:|| ||     | |     |||||||  | |||
Db    282 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTACGGACCAGGA 341

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||  |  |||| || || |||||||| ||| ||||| | |  ||||||
Db    342 GGACCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAGATGTCATTT 401

Qy    517 AACGAAGAAGATGTGTCT 534
          | ||  ||||  || ||
Db    402 GAGGAGAAAGACTTGACT 419


RESULT 4
US-10-164-871-3
; Sequence 3, Application US/10164871
; Patent No. US20020177194A1
; GENERAL INFORMATION:
;  APPLICANT: Hirata, Yuichi
;  TITLE OF INVENTION: STEROID HORMONE BINDING PROTEIN
;  FILE REFERENCE: 06501-059001
;  CURRENT APPLICATION NUMBER: US/10/164,871
;  CURRENT FILING DATE:  2002-06-07
;  PRIOR APPLICATION NUMBER: US/09/565,808
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: WO/JP98/05010
;  PRIOR FILING DATE: 1998-11-06
;  PRIOR APPLICATION NUMBER: JP/9/322376
;  PRIOR FILING DATE: 1997-11-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(669)
US-10-164-871-3

  Query Match             7.9%;  Score 62.2;  DB 9;  Length 672;
  Best Local Similarity   54.6%;  Pred. No. 3.2e-08;
  Matches  171;  Conservative    1;  Mismatches  129;  Indels   12;  Gaps    2;
```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997 05 30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 2776
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-245-78

Query Match 7.9%; Score 62.2; DB 9; Length 2776;
Best Local Similarity 54.6%; Pred. No. 6.5e-08;
Matches 171; Conservative 1; Mismatches 129; Indels 12; Gaps 2;

```
Qy    320 CTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACG 379
          ||   | ||| |||  ||| |||| |   |||||||  |||| ||| ||||| .|  |
Db    314 CTCGCATGAAGAAGCGGGACTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCA 373

Qy    380 AATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAAT 439
          |  |   ||  |||    |||| |||| ||     || |||||:|| ||||  '' |
Db    374 ACCC---GCGCATCCTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGTCAGCA 430

Qy    440 CCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTT 499
            ||||||||| | |  ||    ||    || || || || |  || ||   ||:||
Db    431 AGTTCTACGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGAC 490

Qy    500 TGGGTAAGATGAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTC 550
          |||  |   |  |      |  ||||         |||||  | |  |  |||   || |
Db    491 TGGCCACATTTTGCCTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATT 550

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          | | ||    | | ||  | | | ||   || |||||  |   | |||  |||  :  |||
Db    551 TGAATGCAGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATG 610

Qy    611 CTGTCGTTGGCCG 623
            |   || ||| |
Db    611 ATTATGTAGGCAG 623
```

RESULT 7
US-09-923-876-197
; Sequence 197, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 197
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700156530H1
; NAME/KEY: unsure
; LOCATION: 135
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-197

Query Match 6.9%; Score 54.8; DB 10; Length 253;
Best Local Similarity 68.5%; Pred. No. 2.6e-06;
Matches 74; Conservative 1; Mismatches 33; Indels 0; Gaps 0;

```
Qy    403 ATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGAT 462
          |||||  ||||   | | |||:||||||    | | |   |||||  ||  | ||  ||   |
Db      1 ATCAAGGGCCAGATCTACGACGTCACCCAGAGCAGAATGTTCTATGGACCTGGCGGACCT 60

Qy    463 TACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATG 510
          |||  |  |||||||||| ||||| || |||||||||  | |  ||||||
Db     61 TACGCCCTGTTCGCCGGCAAAGATGCCAGCAGAGCTCTAGCGAAGATG 108
```

RESULT 8
US-10-164-871-1
; Sequence 1, Application US/10164871
; Patent No. US20020177194A1
; GENERAL INFORMATION:
; APPLICANT: Hirata, Yuichi
; TITLE OF INVENTION: STEROID HORMONE BINDING PROTEIN
; FILE REFERENCE: 06501-059001
; CURRENT APPLICATION NUMBER: US/10/164,871
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US/09/565,808
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: WO/JP98/05010
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: JP/9/322376
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(585)
US-10-164-871-1

Query Match 6.9%; Score 54.4; DB 9; Length 588;
Best Local Similarity 54.1%; Pred. No. 5.1e-06;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |'|||| | |   ||||  | |  | | ||||  | | ||  |   ||  ||   |
Db    214 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 270

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||       ||||||||  |||
Db    271 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 330

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   |||  |   | |  |   |  | |
Db    331 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTCTGCCTG 390

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||          |||| ||    |  |||  |  |||||||      ||
Db    391 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 450

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | |||||| | ||| |||||| |  | ||       |||||||  |  ||| |||
```

```
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11743
;    LENGTH: 435
;    TYPE: DNA
;    ORGANISM: Bos taurus
;    OTHER INFORMATION: Clone ID: 50-LIB34-022-Q1-E1-E10
US-09-960-352-11743

  Query Match             5.7%;  Score 44.8;  DB 10;  Length 435;
  Best Local Similarity   52.0%;  Pred. No. 0.0025;
  Matches  128;  Conservative    1;  Mismatches  108;  Indels    9;  Gaps    1;

Qy    387 GCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTA 446
          ||  |||    | ||  |||| ||    || |||||:|| |||| || |     |||||
Db      1 GCGCATCCTGCTTGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCAAGTTCTA 60

Qy    447 CGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAA 506
          |||| | |  ||     ||    || || || | || ||   |||||   |||  |
Db     61 CGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGACTGGCAAC 120

Qy    507 GATGAGTAAGAACGAAG---------AAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGA 557
            |  |    |  |||          ||||||  | |  |  |||   || | | | ||
Db    121 ATTTTGCCTAGATAAAGATGCACTTAAAGATGAATATGATGATCTCTCAGATTTGAATGC 180

Qy    558 GAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGT 617
            | | ||  | | | ||   || |||||  |   | ||| |||  || |||  |   ||
Db    181 TGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATGATTATGT 240

Qy    618 TGGCCG 623
            ||| |
Db    241 AGGCAG 246


RESULT 12
US-09-898-416-5/c
; Sequence 5, Application US/09898416
; Patent No. US20020076759A1
; GENERAL INFORMATION:
;  APPLICANT: Dulac, Catherine
;  APPLICANT:  Axel, Richard
;  TITLE OF INVENTION: Cloning Of Vertebrate Pheromone Receptors And Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 0575/48557A/JPW/ADM
;  CURRENT APPLICATION NUMBER: US/09/898,416
;  CURRENT FILING DATE:  2001-07-03
;  PRIOR APPLICATION NUMBER: 08/731,745
;  PRIOR FILING DATE: 1996-10-18
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;    LENGTH: 1053
;    TYPE: DNA
;    ORGANISM: Rattus sp.
US-09-898-416-5

  Query Match             5.1%;  Score 40.2;  DB 10;  Length 1053;
  Best Local Similarity   53.5%;  Pred. No. 0.083;
  Matches   84;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy    484 GACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTT 543
          || |   ||| |||| ||||  | |  | |  | ||  || |||     |  ||    |||
Db    758 GATGGTGGCATAGCTATGGGACACAATAATTTGGACACAGTAGAACATTGACCCATCCTT 699

Qy    544 GAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCT 603
          ||| |||   ||     |||  ||| | | |  |  ||   |||| | |    ||||||
Db    698 GAATGTCATCCTAGAGTAGAAGACAACATTTTCCAAAATGTAGAGAACCACAAAGAAGCT 639

Qy    604 AAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCT 640
            |  |   |  | ||| | || || | ||| | |||
Db    638 CATGAGAATCATGATGGTCCTGGTGGCCCTTTGCTCT 602


RESULT 13
US-09-878-574-6992
; Sequence 6992, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 6992
;    LENGTH: 252
;    TYPE: DNA
;    ORGANISM: Glycine max
;    OTHER INFORMATION: Clone ID: 701099286H1
US-09-878-574-6992

  Query Match             4.9%;  Score 38.8;  DB 10;  Length 252;
  Best Local Similarity   60.4%;  Pred. No. 0.1;
  Matches   64;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy    137 AACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACATCTCCGGAACA 196
          | | |||  | ||  |||||||  |     |  | || | | | || ||||||   |||
Db      6 ACCAAGGCATATGCAGGTGGAATCATGCCTTAAGATCAGCTCAACACCATCTCAAAAACC 65

Qy    197 ACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGT 242
          | ||| | ||  | |||||  |  || | | | ||||| ||| || |
Db     66 ATGTTGGCTCATTCTCTCAGGCTAACAAGCTTTCTTCTTCTGCTAT 111


RESULT 14
US-09-910-943-363
; Sequence 363, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
;  APPLICANT: Hemmati-Brivanlou, Ali
;  APPLICANT:  Altman, Curtis
;  TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
;  FILE REFERENCE: 7529/1G148US1
;  CURRENT APPLICATION NUMBER: US/09/910,943
;  CURRENT FILING DATE:  2001-07-23
;  NUMBER OF SEQ ID NOS: 742
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 363
;    LENGTH: 755
;    TYPE: DNA
;    ORGANISM: Xenopus laevis
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(755)
;    OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-363

  Query Match             4.9%;  Score 38.6;  DB 10;  Length 755;
  Best Local Similarity   50.6%;  Pred. No. 0.2;
  Matches  123;  Conservative    1;  Mismatches  110;  Indels    9;  Gaps    1;

Qy    390 GATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGG 449
          ||||    | ||  |||| ||    ||||| ||:|| |||   ||||     ||||| ||
Db     83 GATCCTACTGGCTGTCAATGGGAAAGTGTTTGACGTGACCCAGGGAAGCAAATTCTATGG 142

Qy    450 CTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGAT 509
```

```
Db    143 GCCTGATGGCCCGTATGGATTATTTGCTGGTAGAGATGCTTCCAGGGGACTAGCAACGTT 202

Qy    510 GAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAA 560

Db    203 TTGCCTCGATAAAGAGGCGCTCCGAGATGAATATGATGACCTGTCCGACTTAAACGCTGT 262

Qy    561 AGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGG 620

Db    263 ACAGATGGAAAGTGTTCGGGAATGGGAAATGCAGTTTAAAGAAAAGTATGAATATGTTGG 322

Qy    621 CCG 623

Db    323 TCG 325


RESULT 15
US-09-960-352-4582
; Sequence 4582, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4582
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 20-LIB3057-013-Q1-K1-E7
US-09-960-352-4582

  Query Match             4.6%;  Score 36.6;  DB 10;  Length 393;
  Best Local Similarity   53.1%;  Pred. No. 0.55;
  Matches   78;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy    624 TGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAG 683

Db     90 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 149

Qy    684 GATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTA 743

Db    150 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTATTTATTTTATAATTTTT 209

Qy    744 CCATAAGTACCAAATTATCTATGAAAT 770

Db    210 TTATTTTTAAAAAATTATTTATAAAAT 236


Search completed: January  8, 2003, 14:14:26
Job time : 55 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 14:12:25 ; Search time 2177 Seconds
(without alignments)
5869.655 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 7??

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	434.4	55.1	439	10	AV820341	AV820341 AV820341
c	2	409	51.8	436	9	AI996124	AI996124 701550133
c	3	399.6	50.6	400	10	AV536609	AV536609 AV536609
c	4	373.2	47.3	403	14	Z25723	Z25723 ATTS1200 Ve
	5	358.6	45.4	376	10	AV829752	AV829752 AV829752
	6	295.4	37.4	347	14	T42131	T42131 5394 Lambda
c	7	272.8	34.6	420	9	AI995575	AI995575 701674857
	8	244.6	31.0	490	14	BQ791634	BQ791634 E3316 Ch
	9	232.6	29.5	259	10	AV828070	AV828070 AV828070
	10	221.4	28.1	433	9	AA585872	AA585872 28521 Lam
	11	209.8	26.6	429	13	BI417544	BI417544 LjNEST30
	12	198.4	25.1	344	14	BQ791458	BQ791458 E4375 Ch
	13	198.4	25.1	476	10	AW782208	AW782208 sm32h06.
	14	198.4	25.1	498	14	BQ452476	BQ452476 sac87b09
	15	198.4	25.1	512	14	BQ786554	BQ786554 saq59g11
	16	198.4	25.1	518	13	BI470428	BI470428 sah91g10
	17	197.4	25.0	438	13	BI472094	BI472094 sah39d07
	18	196.2	24.8	423	14	BQ296803	BQ296803 san85e04
	19	195.8	24.8	346	10	BE231347	BE231347 00422 le
	20	195.6	24.8	360	10	AW832627	AW832627 sml4h06.
	21	195.6	24.8	367	13	BI095107	BI095107 sae03f04
	22	195.6	24.8	383	10	AW759979	AW759979 sl56e07.
	23	195.6	24.8	411	12	BE802169	BE802169 sr25d01.
	24	195.6	24.8	420	14	BQ079899	BQ079899 san40e06
	25	195.6	24.8	421	13	BM307344	BM307344 sak28a10
	26	195.6	24.8	425	13	BM092206	BM092206 sah09d07
	27	195.6	24.8	442	10	BE611618	BE611618 sq85b08.
c	28	195.2	24.7	444	12	BE822921	BE822921 GM700019
	29	194.8	24.7	357	9	AI440919	AI440919 sa56f12.y
	30	194.8	24.7	406	12	BF067637	BF067637 st67a10.
	31	194.8	24.7	431	10	BE210608	BE210608 so50h03.
	32	194.8	24.7	486	10	AW832150	AW832150 sm20d11.
	33	194.2	24.6	416	10	BE611114	BE611114 sq75h03.
	34	194	24.6	434	13	BM525220	BM525220 sal21c01
	35	193.8	24.6	436	12	BE802709	BE802709 sr44e05.
	36	193.8	24.5	488	12	BG154846	BG154846 sab27g05
	37	192.6	24.4	554	14	BQ623911	BQ623911 USDA-FP_
	38	191.8	24.3	562	12	BF278693	BF278693 GA__Eb00
	39	191.8	24.3	600	9	AI730946	AI730946 BNLGHi822
	40	191.8	24.3	624	12	BG442218	BG442218 GA__Ea00.
	41	191.8	24.3	629	12	BF275349	BF275349 GA__Eb00.
	42	191.8	24.3	643	14	BQ403837	BQ403837 GA__Ed00
	43	191.8	24.3	875	12	BF278692	BF278692 GA__Eb00
	44	191.4	24.3	588	9	AI725454	AI725454 BNLGHi122
	45	191	24.2	304	12	BE802105	BE802105 sr24b10.

ALIGNMENTS

RESULT 1
AV820341/c
LOCUS AV820341 439 bp mRNA linear EST 01-APR-20
DEFINITION AV820341 RAFL11 Arabidopsis thaliana cDNA clone RAFL11-09-L22 3', mRNA sequence.
ACCESSION AV820341
VERSION AV820341.1 GI:19862305
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 439)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M. Ono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: maseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essential as reported previously (Seki et al., 1998). cDNA cleaved with Bam

and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL11-09-L22"
                     /clone_lib="RAFL11"
                     /dev_stage="plants at various developmental stages from
                     germination to mature seeds"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to various
                     treatments (dehydration, cold, high salt, ABA, heat and UV
                     ). Dark-grown plants"
BASE COUNT      134 a    106 c     84 g    115 t
ORIGIN
```

```
  Query Match             55.1%;  Score 434.4;  DB 10;  Length 439;
  Best Local Similarity   99.3%;  Pred. No. 1.3e-85;
  Matches  435;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     351 GCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGG 410
           ||||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db     439 GCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATTTACGTCGCAATCAAAGG 380

Qy     411 CCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGAT 470
           ||||||||||:|||||||||||||||||||||| |||||||||||||||||||||||||
Db     379 CCGTGTGTTCGATGTCACCACCGGAAAATCCTTTTACGGCTCCGGAGGCGATTACTCGAT 320

Qy     471 GTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGT 530
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     319 GTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGT 260

Qy     531 GTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGAC 590
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     259 GTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGAC 200

Qy     591 CAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTTTCTGAG 650
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     199 CAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTTTCTGAG 140

Qy     651 ATTGCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATT 710
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     139 ATTGCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATT 80

Qy     711 TCGTGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAAT 770
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      79 TCGTGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAAT 20

Qy     771 AAATCGGGGATTTCGTGT 788
           ||||||||||||||||||
Db      19 AAATCGGGGATTTCGTGT 2
```

```
RESULT 2
AI996124/c
LOCUS       AI996124                436 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701550133 A. thaliana, Columbia Col-0, inflorescence-2 Arabidopsis
            thaliana cDNA clone 701550133, mRNA sequence.
ACCESSION   AI996124
VERSION     AI996124.1  GI:5843029
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,E., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Erzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            E-mail: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia Col-0"
                     /db_xref="taxon:3702"
                     /clone="701550133"
                     /clone_lib="A. thaliana, Columbia Col-0, inflorescence-2"
                     /tissue_type="inflorescence"
                     /dev_stage="4 - 7 weeks"
                     /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
                     library was derived from untreated inflorescence tissue
                     from Arabidopsis thaliana,  Columbia Col-0, at 4 - 7
                     weeks. Plants were grown in 1:1:1 peat
                     moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C
                     under constant light, and  watered with fertilizer. cDNA
                     synthesis was initiated using a NotI-oligo(dT)  primer.
                     Double-stranded cDNA was blunted, ligated to SalI adaptors
                     , digested with  NotI, size-selected, and cloned into the
                     NotI and SalI sites of the pSPORT vector."
BASE COUNT      128 a    103 c     85 g    118 t      2 others
ORIGIN
```

```
  Query Match             51.8%;  Score 409;  DB 9;  Length 436;
  Best Local Similarity   97.7%;  Pred. No. 5.2e-80;
  Matches  423;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy     327 GAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATAC-AACGGCACCGACGAATCAA 385
           |||||||||||||||||||||||||||||||||||||||    ||| ||||||||||||||
Db     436 GAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATANTNACGTCACCGACGAATCAA 377

Qy     386 AGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCT 445
           |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db     376 AGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGATGTCACCACCGGAAAATCCTTCT 317

Qy     446 ACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTA 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     316 ACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTA 257

Qy     506 AGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGA 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     256 AGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGA 197

Qy     566 TCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTG 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     196 TCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTG 137

Qy     626 TTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGA 685
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136 TTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGA 77

Qy     686 TCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACC 745
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 TCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACC 17

Qy     746 ATAAGTACCAAAT 758
           |||| | || | |
Db      16 ATAATACCAAATT 4
```

RESULT 3

```
Qy    320 CTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACG 379
          ||   | ||| |||  ||| |||| |  ||||||||  |||| ||| ||||| ||  |
Db    287 CTCGCATGAAGAAGCGGGACTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCA 346

Qy    380 AATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAAT 439
          |  |   ||  |||    ||||  |||| ||    || |||||:|| ||||  || |
Db    347 ACCC---GCGCATCCTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCA 403

Qy    440 CCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTT 499
            |||||||||  | |  ||    ||    || || || || |  || ||   |||||
Db    404 AGTTCTACGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGAC 463

Qy    500 TGGGTAAGATGAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTC 550
          |||  |   |  |     |  ||||         |||||  | |  |  |||   || |
Db    464 TGGCCACATTTTGCCTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATT 523

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          | | ||    | | ||  | | | ||   || |||||  |   | |||  |||   || |||
Db    524 TGAATGCAGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATG 583

Qy    611 CTGTCGTTGGCCG 623
            |   || |||  |
Db    584 ATTATGTAGGCAG 596
```

RESULT 5
US-10-098-841-217
; Sequence 217, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 217
; LENGTH: 1936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(752)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1936)
; OTHER INFORMATION: n = a,t,c or g
US-10-098-841-217

Query Match 7.9%; Score 62.2; DB 9; Length 1936;
Best Local Similarity 54.6%; Pred. No. 5.4e-08;
Matches 171; Conservative 1; Mismatches 129; Indels 12; Gaps 2;

```
Qy    320 CTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACG 379
          ||   | ||| |||  ||| |||| |  ||||||||  |||| ||| ||||| ||  |
Db    367 CTCGCATGAAGAAGCGGGACTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCA 426

Qy    380 AATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAAT 439
          |  |   ||  |||    ||||  |||| ||    || |||||:|| ||||  || |
Db    427 ACCC---GCGCATCCTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCA 483

Qy    440 CCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTT 499
            |||||||||  | |  ||    ||    || || || || |  || ||   |||||
Db    484 AGTTCTACGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGAC 543

Qy    500 TGGGTAAGATGAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTC 550
          |||  |   |  |     |  ||||         |||||  | |  |  |||   || |
Db    544 TGGCCACATTTTGCCTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATT 603

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          | | ||    | | ||  | | | ||   || |||||  |   | |||  |||   || |||
Db    604 TGAATGCAGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATG 663

Qy    611 CTGTCGTTGGCCG 623
            |   || |||  |
Db    664 ATTATGTAGGCAG 676
```

RESULT 6
US-09-984-245-78
; Sequence 78, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131

```
Qy    629 TCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTCAGGATCT 688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 TCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGATCT 332

Qy    689 TTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACCATA 748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACCATA 392

Qy    749 AGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 789
          |||||||||||||||||||||||||||||||||||||||
Db    393 AGTACCAAATTATCTATGAAATAAATCGGGGATTTCGTGTT 433


RESULT 11
BI417544
LOCUS       BI417544                429 bp    mRNA    linear   EST 15-AUG-2001
DEFINITION  LjNEST30e2r Lotus japonicus nodule library 5 and 7 week-old Lotus
            japonicus cDNA 5', mRNA sequence.
ACCESSION   BI417544
VERSION     BI417544.1  GI:15188567
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Colebatch,G., Freund,S., Trevaskis,B and Udvardi,M.
  TITLE     Lotus japonicus root nodule ESTs: tools for functional genomics
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Udvardi MK
            Molecular Plant Nutrition
            Max Planck Institute of Molecular Plant Physiology
            Am Muehlenberg 1, 14476 Golm, Germany
            Fax: 49 331 567 8250
            Email: udvardi@mpimp-golm.mpg.de
            Seq primer: T7
            High quality sequence stop: 429.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Lotus japonicus"
                     /cultivar="Gifu (B-129)"
                     /db_xref="taxon:34305"
                     /clone_lib="Lotus japonicus nodule library 5 and 7
                     week-old"
                     /dev_stage="5 and 7 week-old plants"
                     /note="Organ: Nodule; Vector: pSPORT1; Site_1: SalI;
                     Site_2: NotI; The library was prepared using mRNA
                     extracted from nodules of 5 and 7 week-old Lotus plants.
                     Nodules were induced by, and contained Mesorhizobium
                     strain R7A."
BASE COUNT      114 a    118 c    109 g     88 t
ORIGIN

  Query Match             26.6%;  Score 209.8;  DB 13;  Length 429;
  Best Local Similarity   79.4%;  Pred. No. 3.9e-36;
  Matches  247;  Conservative    1;  Mismatches   63;  Indels    0;  Gaps    0;

Qy    325 AAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCA 384
          |||| || ||||||  | |||||| |||| || ||||||||||||||||| |||  |||
Db     79 AAGACAACGATGGAGATGACCGCACAGCAACTGAGCCAATACAACGGCACGGACCCATCG 138

Qy    385 AAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTC 444
          ||||| ||||||||| |  | || ||||| ||||||||:||||||||||||||||||||||
Db    139 AAGCCAATCTACGTCTCGGTGAAGGGCCGCGTGTTCGATGTCACCACCGGAAAATCCTTC 198

Qy    445 TACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGT 504
          |||||| |||| |||| |||| ||||||||||||||| |||||| ||||||||  | |
Db    199 TACGGCCCCGGTGGCGCTTACGCGATGTTCGCCGGAAGAGACGCCAGCAGAGCCCTAGCG 258

Qy    505 AAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAG 564
          !!!!!!!! !!!!!!!! |||||| | || ||    || || || ||| | || |||||||
```

```
Db    259 AAGATGAGTAAGAACGAGGAAGATATCTCCCCCAACCTCGATGGCCTCTCCGACAAAGAG 318

Qy    565 ATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGT 624
          |||      ||| ||||| ||||||| ||| || | |||||||||| ||||| ||||  ||
Db    319 ATCGGCGTTCTCAATGACTGGGAGAACAAGTTCGTAGCTAAGTACCCTGTTGTTGCTCGC 378

Qy    625 GTTGTCTCTTA 635
          |||| |||||
Db    379 CTTGTTTCTTA 389


RESULT 12
BQ791458
LOCUS       BQ791458                344 bp    mRNA    linear   EST 30-JUL-2002
DEFINITION  E4375 Chinese cabbage etiolated seedling library Brassica rapa
            subsp. pekinensis cDNA clone E4375, mRNA sequence.
ACCESSION   BQ791458
VERSION     BQ791458.1  GI:22006420
KEYWORDS    EST.
SOURCE      Brassica rapa subsp. pekinensis.
  ORGANISM  Brassica rapa subsp. pekinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 344)
  AUTHORS   Ryu,S.H., Yang,K.A., Lee,S.Y., Kim,H.-I., Cho,M.J. and Lim,C.O.
  TITLE     Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA
            (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Lim, C.O.
            Plant Molecular Biology & Biotechnology Research Centre
            Gyeongsang National University
            #900 Gazwa-dong, Jinju 660-701, Korea
            Tel: 82 55 751 6255
            Fax: 82 55 759 9363
            Email: colim@nongae.gsnu.ac.kr
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .344
                     /organism="Brassica rapa subsp. pekinensis"
                     /cultivar="Jangwon"
                     /db_xref="taxon:51351"
                     /clone="E4375"
                     /clone_lib="Chinese cabbage etiolated seedling library"
                     /tissue_type="Etiolated seedling"
                     /lab_host="XL-1 Blue"
                     /note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"
BASE COUNT       96 a     97 c     86 g     65 t
ORIGIN

  Query Match             25.1%;  Score 198.4;  DB 14;  Length 344;
  Best Local Similarity   86.8%;  Pred. No. 1.3e-33;
  Matches  217;  Conservative    1;  Mismatches   32;  Indels    0;  Gaps    0;

Qy    329 AAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGC 388
          | |  |||| |||| |||||||||| |||| |||||||| ||||||  ||| ||||
Db     95 AGAGAATGGAGTTCACGGCAGAGCAGCTTAGCCAGTACAACGGGACCGACTCATCGAAGC 154

Qy    389 CGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACG 448
          |||||||||| || |||| |||| || |||||:|||||||||| ||||||||||||||
Db    155 CGATCTACGTGGCGATCAAGGGCCGAGTATTCGACGTCACCACCGGCAAATCCTTCTACG 214

Qy    449 GCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGA 508
          ||   |||||||||||| |||||||||| |||||||||||||||||||| |||| ||||
Db    215 GCGTAGGAGGCGATTACGCGATGTTCGCGGAAAAGACGCGAGCAGAGCTCTGGGGAAGA 274

Qy    509 TGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCA 568
          |||| |||||||||||||| |||||||||||||| ||| ||||||| |||||||||| |
Db    275 TGAGCAAGAACGAAGAAGACGTGTCTCCTTCTCTCGAAAGTCTCACCGAGAAAGAGATGA 334

Qy    569 ATACTCTTAA 578
          ||||||||
Db    335 CCACTCTTAA 344
```

```
RESULT 13
AW782208
LOCUS       AW782208                476 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sm02h06.y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1027-7764 5' similar to TR:O70606 O70606 PUTATIVE PROGESTERONE
            BINDING PROTEIN. ;, mRNA sequence.
ACCESSION   AW782208
VERSION     AW782208.1  GI:7796814
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 476)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Insert Length: 642  Std Error: 0.00
            High quality sequence stop: 406.
FEATURES             Location/Qualifiers
     source          1..476
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1027-7764"
                     /clone_lib="Gm-c1027"
                     /tissue_type="cotyledons of 3- and 7-day-old Williams
                     seedlings"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from cotyledons of 3- and 7-day-old Williams seedlings
                     which were propagated on paper towels with distilled
                     water.  The cotyledons were flash-frozen in liquid
                     nitrogen, then lyophilized for 72 hours.  Unequal amounts
                     of mRNA was used for cDNA synthesis.  Stratagene's cDNA
                     Synthexix Kit (catalog number 200401) was used to
                     synthesize the cDNA.  First- stranded synthesis was
                     performed with 5-methyl dCTP, hence the ligated cDNA was
                     hemimethylated.  A modification of Stratagene's
                     first-strand  synthesis primer was used.  An anchor
                     nucleotide (V=A, C, or G) was added to the 3' end of the
                     primer [GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18] to anchor
                     the primer at the 5' end of the poly(A) tract.  After
                     second- strand synthesis, the cDNA ends were filled in
                     with cloned Pfu DNA, ligated to EcoRI adapters and
                     subsequently phosphorylated.  The XhoI site within the
                     first-strand synthesis primer was then restricted by
                     digestion with XhoI; all XhoI sites in the cDNA would be
                     protected by their hemimethylated status.  The cDNA
                     constructs were size-fractionated with a 500 bp cutoff,
                     using GibcoBRL Life Technologies' cDNA Size Fractionation
                     column.  The column eluent was then ligated into
                     Stratagene's pBluescript(tm) II XR Predigested vector
                     (pBluescript II SK(+) that has been digested with EcoRI
                     and XhoI, and phosphorylated by Stratagene).  97% of the
                     white and blue colonies appear to contain recombinant
                     plasmids with cDNA inserts, based on size (n=30).  This
                     library was constructed by Dr. Paul Keim and Dr. Virginia
                     Coryell."
BASE COUNT      123 a    144 c    103 g    106 t
ORIGIN
```

```
  Query Match             25.1%;  Score 198.4;  DB 10;  Length 476;
  Best Local Similarity   68.2%;  Pred. No. 1.2e-33;
  Matches  288;  Conservative    1;  Mismatches  132;  Indels    1;  Gaps    1;

Qy    328 AAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAG 387
          || |  |||||  | ||  |  ||||||| ||||||||||||||||||| |||  ||| |||
Db     53 AAGACAATGGAGCTGACACCCCAGCAGCTGAGCCAATACAACGGCACGGACCCATCGAAG 112

Qy    388 CCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTAC 447
          |||||||||||| ||  | || ||||| || | |||:||||||||||||||||||||| |||
Db    113 CCGATCTACGTGGCGGTGAAGGGCCGCGTCTACGACGTCACCACCGGAAAATCCTTTTAC 172

Qy    448 GGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAG 507
          ||| |||| |||   ||| | |||||||||||| |||||||| |||||||| |||  |||
Db    173 GGCCCCGGCGGCCCCTACGCCATGTTCGCCGGCAAAGACGCCAGCAGAGCCCTGGCGAAG 232

Qy    508 ATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          ||||| |||||||| || ||  | || || || || || || ||| | || || ||||||
Db    233 ATGAGCAAGAACGACGACGACATCTCCCCCTCCCTCGACGGCCTCTCCGACAAGGAGATC 292

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTT 627
              ||| || || ||||||| ||||||  |||||||||| |||||||||| ||||||
Db    293 GGCGTTCTCAACGACTGGGAGAACAAATTCCAAGCTAAGTACCCTGTCGTTGCTCGTGTT 352

Qy    628 GTCTCTTAGGTCTCT-CTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGAT 686
            ||  |||    | |  |       | || | | |||  | || |   || |||  | |||
Db    353 CTCAATTAATCATATGATAACCTAAGTGTATTTTGTACTCTATCATGTGAGTGGCATGAT 412

Qy    687 CTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACCA 746
          || | |T|              ||  |  ||    || ||  | || ||| |||| |
Db    413 CTCTTTCTGTAACCGAACTCAATCACCCTTTAATCCCTAATGATATTAATAAAATTCCTC 472

Qy    747 TA 748
          ||
Db    473 TA 474
```

```
RESULT 14
BQ452476
LOCUS       BQ452476                498 bp    mRNA    linear   EST 29-MAY-2002
DEFINITION  sao87b09.y1 Gm-c1081 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1081-2753 5' similar to TR:Q9SK39 Q9SK39 PUTATIVE STEROID
            BINDING PROTEIN. ;, mRNA sequence.
ACCESSION   BQ452476
VERSION     BQ452476.1  GI:21255588
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
```

Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Putative full length read
vector to vector length is 556
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES Location/Qualifiers
source 1. .498
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1081-2753"
/clone_lib="Gm-c1081"
/tissue_type="Roots of 7 day old 'Bragg' seedlings"
/dev_stage="7 days old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of 7 day old 'Bragg' seedlings that were mock-infected 48 hours prior to harvest. Dr. Gary Stacey generously donated the tissue. The roots were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first- strand synthesis primer was then restricted by digestion with XhoI from Promega (400/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl S-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constucted in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University."

BASE COUNT 135 a 147 c 101 g 115 t
ORIGIN

Query Match 25.1%; Score 198.4; DB 14; Length 498;
Best Local Similarity 68.2%; Pred. No. 1.2e-33;
Matches 288; Conservative 1; Mismatches 132; Indels 1; Gaps 1;

```
Qy    328 AAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAG 387
          || |  |||||  | ||  |  ||||||| |||||||||||||||||| |||  ||| |||
Db     36 AAGACAATGGAGCTGACACCCCAGCAGCTGAGCCAATACAACGGCACGGACCCATCGAAG 95

Qy    388 CCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTAC 447
          |||||||||| ||  | || ||||| || | |||:||||||||||||||||||| |||
Db     96 CCGATCTACGTGGCGGTGAAGGGCCGCGTCTACGACGTCACCACCGGAAAATCCTTTTAC 155

Qy    448 GGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAG 507
          ||| |||| |||   ||| | |||||||||||| |||||||| ||||||||  |||  |||
Db    156 GGCCCCGGCGGCCCCTACGCCATGTTCGCCGGCAAAGACGCCAGCAGAGCCCTGGCGAAG 215

Qy    508 ATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          ||||| |||||||| || ||  | || || || || || || ||| | || || ||||||
Db    216 ATGAGCAAGAACGACGACGACATCTCCCCCTCCCTCGACGGCCTCTCCGACAAGGAGATC 275
```

```
Qy    568 AATACTCTTA..GATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTT 627
          ||| |  || ||||||| ||||||  |||||||||| |||||||||  ||||||
Db    276 GGCGTTCTCA..GACTGGGAGAACAAATTCCAAGCTAAGTACCCTGTCGTTGCTCGTGTT 335

Qy    628 GTCTCTTAGG..TCT-CTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGAT 686
          ||  |||    | |  |      | || | | |||  | || |   || |||  | |||
Db    336 CTCAATTAAT.ATATGATAACCTAAGTGTATTTTGTACTCTATCATGTGAGTGGCATGAT 395

Qy    687 CTTTGTGTTG..TGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACCA 746
          || | | |            || |   ||     || ||  | || ||| |||| |
Db    396 CTCTTTCTG..ACCGAACTCAATCACCCTTTAATCCCTAATAATATTAATAAAATTCCTC 455

Qy    747 TA 748
          ||
Db    456 TA 457
```

RESULT 15
BQ786554
LOCUS BQ786554 512 bp mRNA linear EST 26-JUL-2002
DEFINITION saq69g11.y1 Gm-c1076 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1076-5301 5' similar to TR:Q9SK39 Q9SK39 PUTATIVE STEROID BINDING PROTEIN. ;, mRNA sequence.
ACCESSION BQ786554
VERSION BQ786554.1 GI:21995026
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 512)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolon,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Watterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 421.

FEATURES Location/Qualifiers
source 1. .512
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1076-5301"
/clone_lib="Gm-c1076"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI

adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodkin lab , University of Illinois)."

B. E COUNT 134 a 153 c 110 g 115 t
C TION

Query Match 25.1%; Score 198.4; DB 14; Length 512;
Best Local Similarity 68.2%; Pred. No. 1.2e-33;
Matches 288; Conservative 1; Mismatches 132; Indels 1; Gaps 1;

```
Q      328 AAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAG 387
           || |  |||||  | ||  |  ||||||| ||||||||||||||||| |||  ||| |||
D       61 AAGACAATGGAGCTGACACCCCAGCAGCTGAGCCAATACAACGGCACGGACCCATCGAAG 120

Q      388 CCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTAC 447
           |||||||||||  ||  | || ||||| || | |||:||||||||||||||||||| |||
D      121 CCGATCTACGTGGCGGTGAAGGGCCGCGTCTACGACGTCACCACCGGAAAATCCTTTTAC 180

Q      448 GGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAG 507
           ||| |||| |||   ||| | |||||||||||| |||||||| |||||||  |||  |||
D      181 GGCCCCGGCGGCCCCTACGCCATGTTCGCCGGCAAAGACGCCAGCAGAGCCCTGGCGAAG 240

Q      508 ATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
           ||||| |||||||| || ||  | || || || || || || ||| | || || ||||||
D      241 ATGAGCAAGAACGACGACGACATCTCCCCCTCCCTCGACGGCCTCTCCGACAAGGAGATC 300

Q      568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTT 627
           |||  || || |||||||  ||||||  |||||||||| ||||||||||  ||||||
D      301 GGCGTTCTCAACGACTGGGAGAACAAATTCCAAGCTAAGTACCCTGTCGTTGCTCGTGTT 360

Q      628 GTCTCTTAGGTCTCT-CTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGAT 686
           ||  |||     | |  |     | || | | |||  | || |   || |||  | |||
D      361 CTCAATTAATCATATGATAACCTAAGTGTATTTTGTACTCTATCATGTGAGTGGCATGAT 420

Q      687 CTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTACCA 746
           || | | |              ||  |   ||     || ||  | || ||| |||| |
D      421 CTCTTTCTGTAACCGAACTCAATCACCCTTTAATCCCTAATAATATTAATAAAATTCCTC 480

Q      747 TA 748
           ||
D      481 TA 482
```

Search completed: January 8, 2003, 17:44:39
Job time : 2194 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 10:55:28 ; Search time 260 Seconds
(without alignments)
6833.949 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	788.6	99.9	789	21	AAC52188	Arabidopsis thalia
2	489.2	62.0	522	21	AAC35978	Arabidopsis thalia
3	402.2	51.0	474	21	AAC32918	Arabidopsis thalia
4	282.8	35.8	477	21	AAC34933	Arabidopsis thalia
5	276.8	35.1	413	21	AAC36826	Arabidopsis thalia
6	115.4	14.6	437	21	AAC40417	Arabidopsis thalia
7	112.2	14.2	455	21	AAC50186	Arabidopsis thalia
8	108.4	13.7	252	21	AAC50177	Arabidopsis thalia
9	100.2	12.7	506	21	AAC44432	Arabidopsis thalia
10	91.2	11.[illegible]	655	21	AAZ61724	cDNA encoding rat
11	91.2	11.[illegible]	655	22	AAC99617	Skin cell cDNA, SE
12	91.2	11.[illegible]	655	24	ABL34809	Rat cDNA isolated
13	89.6	11.4	751	24	ABQ65424	Arabidopsis thalia
14	89.6	11.4	887	21	AAC50344	Arabidopsis thalia
15	89.6	11.4	889	21	AAC34295	Arabidopsis thalia
16	85	10.8	622	22	AAH87745	Peppermint plant o
17	75.8	9.6	516	21	AAA15924	Human protein clon
18	75.8	9.6	673	21	AAA15934	Human protein clon
19	75.8	9.6	770	21	AAC47068	Arabidopsis thalia
20	74	9.4	575	22	AAH87744	Peppermint plant o
21	67	8.5	2127	21	AAC52139	Arabidopsis thalia
22	63.2	8.[illegible]	825	21	AAC45975	Arabidopsis thalia
23	63.2	8.[illegible]	913	21	AAC45378	Arabidopsis thalia
24	62.2	7.9	672	20	AAV72313	Human hSMBP2 codin
25	62.2	7.9	1081	24	ABL95727	Human angiogenesis
26	62.2	7.9	1081	24	ABL88238	Human PRO6008 cDNA
27	62.2	7.9	1936	22	AAI58011	Human polynucleoti
28	62.2	7.9	2776	20	AAX00678	Human secreted pro
C 29	62.2	7.9	2833	22	AAI59797	Human polynucleoti
30	61.6	7.8	919	21	AAC35670	Arabidopsis thalia
31	61	7.7	643	21	AAF13129	Aspergillus oryzae
C 32	60.4	7.7	789	21	AAC52188	Arabidopsis thalia
33	57.2	7.2	437	21	AAF08541	Fusarium venenatum
34	56.4	7.1	1000	23	ABL14777	Drosophila melanog
35	56	7.1	1893	19	AAV09852	Pig plasma membran
36	55.2	7.0	1868	19	AAV63170	cDNA from clone bp
37	54.4	6.9	588	20	AAV72312	Human hSMBP1 codin
38	54.4	6.9	788	19	AAV60245	Nucleotide sequenc
39	54.4	6.9	1254	20	AAZ42132	Human normal bladd
40	54.4	6.9	1875	20	AAV84368	Human stomach carc
41	54.4	6.9	1890	24	ABN97363	Gene #3861 used to
42	54.4	6.9	1890	24	ABK64836	Human benign prost
43	54.4	6.9	1890	24	ABL68142	Kidney cancer rela
44	54.4	6.9	1890	24	ABL68368	Kidney cancer rela
45	54.4	6.9	1918	22	AAH57484	Human liver cell s

ALIGNMENTS

RESULT 1
AAC52188
ID AAC52188 standard; DNA; 789 BP.
XX
AC AAC52188;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70554.
XX
KW Hybridisation assay; genetic mapping, gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US 0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             99.9%;  Score 788.6;  DB 21;  Length 789;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-226;
  Matches  789;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTT 300

Qy     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGC 360

Qy     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTC 420

Qy     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGA 480

Qy     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCT 540

Qy     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAA 600

Qy     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTAT 660

Qy     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGG 720

Qy     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGA 780

Qy     781 TTTCGTGTT 789
          |||||||||
Db     781 TTTCGTGTT 789


RESULT 2
AAC35978
ID   AAC35978 standard; DNA; 522 BP.
XX
AC   AAC35978;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 12108.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
```

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.0%; Score 489.2; DB 21; Length 522;
Best Local Similarity 99.0%; Pred. No. 3e-136;
Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy    294 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
          |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA  73

Qy    354 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy    414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
          |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 193

Qy    474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 253

Qy    534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 313

Qy    594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 373

Qy    654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 433

Qy    714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 493

Qy    774 TCGGGGATTTCGTGTT 789
          ||||||||||||||||
Db    494 TCGGGGATTTCGTGTT 509
```

RESULT 3
AAC32918
ID AAC32918 standard; DNA; 474 BP.
XX
AC AAC32918;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1123.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.

```
P   20-JUL-1999;   99US-0144884.
P   21-JUL-1999;   99US-0144814.
P   21-JUL-1999;   99US-0145086.
P   21-JUL-1999;   99US-0145088.
P   22-JUL-1999;   99US-0145085.
P   22-JUL-1999;   99US-0145087.
P   22-JUL-1999;   99US-0145089.
P   22-JUL-1999;   99US-0145192.
P   23-JUL-1999;   99US-0145145.
P   23-JUL-1999;   99US-0145218.
P   23-JUL-1999;   99US-0145224.
P   26-JUL-1999;   99US-0145276.
P   27-JUL-1999;   99US-0145913.
P   27-JUL-1999;   99US-0145918.
P   27-JUL-1999;   99US-0145919.
P   28-JUL-1999;   99US-0145951.
P   02-AUG-1999;   99US-0146386.
P   02-AUG-1999;   99US-0146388.
P   02-AUG-1999;   99US-0146389.
P   03-AUG-1999;   99US-0147038.
P   04-AUG-1999;   99US-0147204.
P   04-AUG-1999;   99US-0147302.
P   05-AUG-1999;   99US-0147192.
P   05-AUG-1999;   99US-0147260.
P   06-AUG-1999;   99US-0147303.
P   06-AUG-1999;   99US-0147416.
P   09-AUG-1999;   99US-0147493.
P   09-AUG-1999;   99US-0147935.
P   10-AUG-1999;   99US-0148171.
P   11-AUG-1999;   99US-0148319.
P   12-AUG-1999;   99US-0148341.
P   13-AUG-1999;   99US-0148565.
P   13-AUG-1999;   99US-0148684.
P   16-AUG-1999;   99US-0149368.
P   17-AUG-1999;   99US-0149175.
P   18-AUG-1999;   99US-0149426.
P   20-AUG-1999;   99US-0149722.
P   20-AUG-1999;   99US-0149723.
P   20-AUG-1999;   99US-0149929.
P   23-AUG-1999;   99US-0149902.
P   23-AUG-1999;   99US-0149930.
P   25-AUG-1999;   99US-0150566.
P   26-AUG-1999;   99US-0150884.
P   27-AUG-1999;   99US-0151065.
P   27-AUG-1999;   99US-0151066.
P   27-AUG-1999;   99US-0151080.
P   30-AUG-1999;   99US-0151303.
P   31-AUG-1999;   99US-0151438.
P   01-SEP-1999;   99US-0151930.
P   07-SEP-1999;   99US-0152363.
P   10-SEP-1999;   99US-0153070.
P   13-SEP-1999;   99US-0153758.
P   15-SEP-1999;   99US-0154018.
P   16-SEP-1999;   99US-0154039.
P   20-SEP-1999;   99US-0154779.
P   22-SEP-1999;   99US-0155139.
P   23-SEP-1999;   99US-0155486.
P   24-SEP-1999;   99US-0155659.
P   28-SEP-1999;   99US-0156458.
P   29-SEP-1999;   99US-0156596.
P   04-OCT-1999;   99US-0157117.
P   05-OCT-1999;   99US-0157753.
P   06-OCT-1999;   99US-0157865.
P   07-OCT-1999;   99US-0158029.
P   08-OCT-1999;   99US-0158232.
P   12-OCT-1999;   99US-0158369.
P   13-OCT-1999;   99US-0159293.
P   13-OCT-1999;   99US-0159294.
P   13-OCT-1999;   99US-0159295.
P   14-OCT-1999;   99US-0159329.
P   14-OCT-1999;   99US-0159330.
P   14-OCT-1999;   99US-0159331.
P   14-OCT-1999;   99US-0159637.
PR  14-OCT-1999;   99US-0159638.
PR  18-OCT-1999;   99US-0159584.
PR  21-OCT-1999;   99US-0160741.
PR  21 OCT 1999;   99US 0160767.
PR  21-OCT-1999;   99US-0160768.
PR  21-OCT-1999;   99US-0160770.
PR  21-OCT-1999;   99US-0160814.
PR  21-OCT-1999;   99US-0160815.
PR  22-OCT-1999;   99US-0160980.
PR  22-OCT-1999;   99US-0160981.
PR  22-OCT-1999;   99US-0160989.
PR  25-OCT-1999;   99US-0161404.
PR  25-OCT-1999;   99US-0161405.
PR  25-OCT-1999;   99US-0161406.
PR  26-OCT-1999;   99US-0161359.
PR  26-OCT-1999;   99US-0161360.
PR  26-OCT-1999;   99US-0161361.
PR  28-OCT-1999;   99US-0161920.
PR  28-OCT-1999;   99US-0161992.
PR  28-OCT-1999;   99US-0161993.
PR  29-OCT-1999;   99US-0162142.

  Query Match             51.0%;  Score 402.2;  DB 21;  Length 474;
  Best Local Similarity   98.8%;  Pred. No. 3.4e-110;
  Matches  404;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    295 TCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAG 354
          ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAG 60

Qy    355 CTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGT 414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGT 120

Qy    415 GTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTC 474
          |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTC 180

Qy    475 GCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCT 534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCT 240

Qy    535 CCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAA 594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAA 300

Qy    595 TTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTG 654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTG 360

Qy    655 CACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTT 703
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTT 409


RESULT 4
AAC34933
ID   AAC34933 standard; DNA; 477 BP.
XX
AC   AAC34933;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 8419.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
```

PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 35.8%; Score 282.8; DB 21; Length 477;
Best Local Similarity 99.3%; Pred. No. 2.1e-74;
Matches 284; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 62

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 122

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 182

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 242

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGA 286
           |||||||||||||||||||||||||||||||||||| | ||||
Db     243 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 288
```

RESULT 5
AAC36826
ID AAC36826 standard; DNA; 413 BP.
XX
AC AAC36826;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15206.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147116.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 35.1%; Score 276.8; DB 21; Length 413;
Best Local Similarity 99.3%; Pred. No. 1.2e-72;
Matches 278; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      7 AACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTCTGAAGA 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTCTGAAGA 60

Qy     67 AAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGAGGCA 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGAGGCA 120

Qy    127 TTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACAT 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACAT 180

Qy    187 CTCCGGAACAAAGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGTCTCC 246
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTCCGGAACAAAGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGTCTCC 240

Qy    247 GCAGCCGTTAC TCCTCTGGTGAGAGCGAGAAGACGAAGA 286
          |||||||||| |||||||||||||||||| | ||||
Db    241 GCAGCCGTTACGTCCTCTGGTGAGAGCGAGAAGGCTAAGA 280
```

ULT 6
40417
AAC40417 standard; DNA; 437 BP.

AAC40417;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 28228.

Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999; 99US-0129845.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

```
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match          14.6%;  Score 115.4;  DB 21;  Length 437;
  Best Local Similarity  77.3%;  Pred. No. 2.9e-24;
  Matches  140;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy      63 AAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGA 122
           |||| | ||||| ||| |||||||| ||||||  || || |  |||||||||||||| ||
Db      50 AAGAGAAATGAGCTCTGCAAGCAAAACATGGATGGTTGCTGCAAGCATCGGAGCCGTTGA 109

Qy     123 GGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCA 182
           |||||| ||||||||||||||  | ||||| |||||||||  |  |||| || |  |||||
Db     110 GGCATTGAAAGACCAACTAGGGGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAATCA 169

Qy     183 ACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGT 242
             ||||||| |||||| | ||||| || || ||||  || | | ||||||| || ||  |
Db     170 GTATCTCCGCAACAACTTAAGATCCGTGTCGCAAGCTAAGACGCTCTCTTCGTCATCATT 229

Qy     243 C 243
           |
Db     230 C 230


RESULT 7
AAC50186
ID   AAC50186 standard; DNA; 455 BP.
XX
AC   AAC50186;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 63895.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
```

18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 14.2%; Score 112.2; DB 21; Length 455;
Best Local Similarity 76.2%; Pred. No. 2.7e-23;
Matches 138; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy      63 AAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGA 122
           |||| | ||||| ||| |||||||| | ||||  || || |  ||||||||||||||| ||
Db      50 AAGAGAAATGAGCTCTGCAAGCAAAACGTGGATGGTTGCTGCAAGCATCGGAGCCGTTGA 109

Qy     123 GGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCA 182
           |||||| ||||||||||||||  | ||||| ||||||||| |  |||| || |  |||||
Db     110 GGCATTGAAAGACCAACTAGGCGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAATCA 169
```

```
Qy     183 ACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGT 242
           |||| || |||||| | ||||| || || ||||  || | | ||||||| || ||  |
Db     170 GTATCTACGCAACAACTTAAGATCCGTGTCGCAAGCTAAGAAGCTCTCTTCCTCATCAAT 229

Qy     243 C 243
           |
Db     230 C 230


RESULT 8
AAC50177
ID   AAC50177 standard; DNA; 252 BP.
XX
AC   AAC50177;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 63866.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138510.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
```

10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.

Query Match 13.7%; Score 108.4; DB 21; Length 252;
Best Local Similarity 76.4%; Pred. No. 2.8e-22;
Matches 133; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy      70 ATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGAGGCATTA 129
           ||||| ||| |||||||| | ||||  || || |  ||||||||||||| ||||||||
Db       1 ATGAGCTCTGCAAGCAAAACGTGGATGGTTGCTGCAAGCATCGGAGCCGTTGAGGCATTG 60

Qy     130 AAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACATCTC 189
           ||||||||||||||  | ||||| |||||||||| |  |||| || |  |||||  ||||
Db      61 AAAGACCAACTAGGCGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAATCAGTATCTA 120

Qy     190 CGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCTGTC 243
           || |||||| | ||||| || || ||||  || | | ||||||| || ||  ||
Db     121 CGCAACAACTTAAGATCCGTGTCGCAAGCTAAGAAGCTCTCTTCCTCATCAATC 174
```

RESULT 9
AAC44432
ID AAC44432 standard; DNA; 506 BP.
XX
AC AAC44432;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 42806.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 21-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 12.7%; Score 100.2; DB 21; Length 506;
Best Local Similarity 97.1%; Pred. No. 1.1e-19;
Matches 102; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 297 CCTTGAGAAA AATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCT 356

```
Db    402 CCGCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCT 461

Qy    357 AAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGC 401
          |||||||||||||||||||||||||||||||||||||||||||||
Db    462 AAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGC 506


RESULT 10
AAZ61724
ID   AAZ61724 standard; cDNA; 655 BP.
XX
AC   AAZ61724;
XX
DT   27-MAR-2000  (first entry)
XX
DE   cDNA encoding rat dermal papilla protein DP3, SEQ ID NO:119.
XX
KW   Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW   embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW   secreted; transmembrane; inflammation; cancer; neurological disease;
KW   angiogenesis; tumour vascularisation; growth disorder;
KW   developmental disorder; skin wound; hair follicle disorder;
KW   anti-inflammatory; cytostatic; neuroprotective; vulnery; ss.
XX
OS   Rattus sp.
XX
PN   WO9955865-A1.
XX
PD   04-NOV-1999.
XX
PF   29-APR-1999;   99WO-NZ00051.
XX
PR   29-APR-1998;   98US-0069726.
PR   09-NOV-1998;   98US-0188930.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Strachan L,  Sleeman M,  Watson JD,  Onrust R,  Kumble A,  Murison JG;
XX
DR   WPI; 2000-072177/06.
DR   P-PSDB; AAY76019.
XX
PT   Novel polynucleotides useful for the treatment of various conditions
PT   including wounds and cancer -
XX
PS   Claim 1; Page 100; 235pp; English.
XX
CC   The invention relates to novel nucleic acid sequences derived from rat
CC   dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC   and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC   cells. Polypeptides of the invention may be used to treat inflammation,
CC   cancer and neurological diseases. The proteins may be used to stimulate
CC   the growth and motility of keratinocytes, to inhibit the growth of
CC   cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC   modulate skin inflammation, to modulate epithelial cell growth and to
CC   inhibit binding of HIV-1 to leukocytes. The invention may also be used
CC   to treat growth and developmental defects, skin wounds and hair follicle
CC   disorders. Sequences AAZ61606-Z61832 represent cDNA sequences derived
CC   from several mouse, rat or human skin cell types. Sequences
CC   AAZ61606-Z61649, AAZ61725-Z61765, AAZ61802-Z61811 and AAZ61826 encode
CC   proteins with an N-terminal signal sequence, indicating that the proteins
CC   are secreted. Sequences AAZ61650-Z61668, AAZ61766-Z61780, AAZ61812-Z61817
CC   and AAZ61827-Z61829 encode proteins with one or more putative
CC   transmembrane domains.
XX
SQ   Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 other;

  Query Match             11.6%;  Score 91.2;  DB 21;  Length 655;
  Best Local Similarity   60.7%;  Pred. No. 6.4e-17;
  Matches  147;  Conservative    1;  Mismatches   94;  Indels    0;  Gaps    0;

Qy    340 TTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTC 399
          ||||||| ||| ||||  ||  |||| ||||  || ||     | || |||||| |
Db    136 TTCACCGAGGAGGAGCTGGCCCGCTACAGCGGCGAGGAGGAGGATCAACCCATCTACTTG 195

Qy    400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
          ||| | || ||    ||||||||:|||||| | || ||    || || ||    |||| |
Db    196 GCAGTGAAGGGAGTGGTGTTCGATGTCACCTCTGGGAAGGAGTTTTATGGACGTGGAGCC 255

Qy    460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
          |||       || ||||| || ||| ||||||||| | |||  ||||||     | |
Db    256 CCCTACAACGCCTTGGCCGGGAAGGACTCGAGCAGAGGTGTGGCCAAGATGTCGCTGGAT 315

Qy    520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAAT 579
          | |||  |  ||| |    ||   |||||||||  || ||| |  |  | ||  ||
Db    316 CCTGCAGACCTCACTCATGACATTTCTGGTCTCACTGCCAAGGAGCTGGAAGCCCTCGAT 375

Qy    580 GA 581
          ||
Db    376 GA 377


RESULT 11
AAC99657
ID   AAC99657 standard; cDNA; 655 BP.
XX
AC   AAC99657;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Skin cell cDNA, SEQ ID NO: 119.
XX
KW   Rat; skin cell; cytostatic; antiinflammatory; anti-HIV;
KW   nootropic; neuroprotective; vulnerary; immunomodulatory; vaccine;
KW   keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW   inflammation; neurological disease; ss.
XX
OS   Rattus sp.
XX
PN   WO200069894-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-NZ00075.
XX
PR   14-MAY-1999;   99US-0312283.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Onrust R,  Sleeman M,  Kumble KD,  Murison JG;
XX
DR   WPI; 2001-007495/01.
DR   P-PSDB; AAB55958.
XX
PT   New isolated polynucleotide used in the identification of genetic
PT   disorders and encoding polypeptides used for treating inflammatory
PT   disease, cancer and neurological diseases -
XX
PS   Claim 1; Page 133-134; 352pp; English.
XX
CC   The present polynucleotide encodes a polypeptide which is expressed in
CC   mammalian skin cells. The polypeptide is useful for stimulating
CC   keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC   modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC   tumours, modulating skin inflammation, stimulating the growth of
CC   epithelial cells, inhibiting the binding of human immunodeficiency virus
CC   (HIV)-1 to leukocytes, and treating inflammmatory disease, cancer and
CC   neurological diseases. The polynucleotide can be used as a marker, in
CC   the identification of genetic disorders, and for the design of
CC   oligonucleotides for examining expression patterns.
XX
SQ   Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 other;

  Query Match             11.6%;  Score 91.2;  DB 22;  Length 655;
  Best Local Similarity   60.7%;  Pred. No. 6.4e-17;
```

```
  Matches  147;  Conservative    1;  Mismatches   94;  Indels    0;  Gaps    0;

Qy    340 TTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTC 399
          ||||||| ||| ||||  ||  |||| ||||  || ||     | || ||||||
Db    136 TTCACCGAGGAGGAGCTGGCCCGCTACAGCGGCGAGGAGGAGGATCAACCCATCTACTTG 195

Qy    400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
          ||| | || ||    ||||||||:|||||| | || ||    || || ||    |||| |
Db    196 GCAGTGAAGGGAGTGGTGTTCGATGTCACCTCTGGGAAGGAGTTTTATGGACGTGGAGCC 255

Qy    460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
            |||       || ||||| || ||| |||||||||| | |||  ||||||     | |
Db    256 CCCTACAACGCCTTGGCCGGGAAGGACTCGAGCAGAGGTGTGGCCAAGATGTCGCTGGAT 315

Qy    520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAAT 579
          | |||  |  ||| |    ||  ||||||||||  || ||| |  |  | ||  ||
Db    316 CCTGCAGACCTCACTCATGACATTTCTGGTCTCACTGCCAAGGAGCTGGAAGCCCTCGAT 375

Qy    580 GA 581
          ||
Db    376 GA 377


RESULT 12
ABL34809
ID   ABL34809 standard; cDNA; 655 BP.
XX
AC   ABL34809;
XX
DT   04-APR-2002  (first entry)
XX
DE   Rat cDNA isolated from skin cells SEQ ID NO: 119.
XX
KW   Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW   developmental defect; inflammatory disease; dermatological; vulnerary;
KW   immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;
KW   ss.
XX
OS   Rattus sp.
XX
PN   WO200190357-A1.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-NZ00099.
XX
PR   24-MAY-2000; 2000US-206650P.
PR   25-JUL-2000; 2000US-221232P.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Strachan L,  Sleeman M,  Onrust R,  Murison JG,  Kumble KD;
XX
DR   WPI; 2002-122020/16.
XX
PT   New polynucleotides and polypeptides encoded by the polynucleotides
PT   isolated from skin cells, useful for treating skin wounds, cancers,
PT   growth and developmental defects, inflammatory diseases, or for
PT   modulating immune responses -
XX
PS   Claim 1; Page 116; 466pp; English.
XX
CC   The present invention provides the protein and coding sequences of cDNAs
CC   isolated from human, murine and rat skin cell libraries. The sequences
CC   can be used in the development of therapeutic agents useful in the
CC   treatment of skin diseases, including skin wounds, cancer, growth
CC   defects, developmental defects and inflammatory diseases. The proteins
CC   have important roles in the induction of hair growth, cell proliferation
CC   and cell-cell interaction, in maintaining tissue integrity, in wound
CC   healing and in modulating immune responses. The present sequence is a
CC   cDNA of the invention.
XX
SQ   Sequence 655 BP; 155 A; 180 C; 203 G; 117 T; 0 other;
```

```
  Query Match           11.6%;  Score 91.2;  DB 24;  Length 655;
  Best Local Similarity  60.7%;  Pred. No. 6.4e-17;
  Matches  147;  Conservative    1;  Mismatches   94;  Indels    0;  Gaps    0;

Qy    340 TTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTC 399
          ||||||| ||| ||||  ||  |||| ||||  || ||     | || ||||||
Db    136 TTCACCGAGGAGGAGCTGGCCCGCTACAGCGGCGAGGAGGAGGATCAACCCATCTACTTG 195

Qy    400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
          ||| | || ||    ||||||||:|||||| | || ||    || || ||    |||| |
Db    196 GCAGTGAAGGGAGTGGTGTTCGATGTCACCTCTGGGAAGGAGTTTTATGGACGTGGAGCC 255

Qy    460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
            |||       || ||||| || ||| |||||||||| | |||  ||||||     | |
Db    256 CCCTACAACGCCTTGGCCGGGAAGGACTCGAGCAGAGGTGTGGCCAAGATGTCGCTGGAT 315

Qy    520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAAT 579
          | |||  |  ||| |    ||  ||||||||||  || ||| |  |  | ||  ||
Db    316 CCTGCAGACCTCACTCATGACATTTCTGGTCTCACTGCCAAGGAGCTGGAAGCCCTCGAT 375

Qy    580 GA 581
          ||
Db    376 GA 377


RESULT 13
ABQ65424
ID   ABQ65424 standard; DNA; 751 BP.
XX
AC   ABQ65424;
XX
DT   21-AUG-2002  (first entry)
XX
DE   Arabidopsis thaliana polynucleotide SEQ ID NO 1.
XX
KW   Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;
KW   stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;
KW   insecticide; antibiotic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   US2002059663-A1.
XX
PD   16-MAY-2002.
XX
PF   26-JAN-2001; 2001US-0770149.
XX
PR   27-JAN-2000; 2000US-178506P.
XX
PA   (GORL/) GORLACH J.
PA   (ANYY/) AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES T M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
PI   Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
PI   Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
```

PI Hurban P;
XX
DR WPI; 2000-479224/51.
XX
PT New nucleic acid that hybridizes to Arabidopsis thaliana sequences,
PT useful e.g. for preparing transgenic plants with increased resistance
PT or altered metabolism -
XX
PS Claim 1; SEQ ID NO 1; 40pp + Sequence Listing; English.
XX
CC The invention relates to nucleic acids (I) that hybridise under stringent
CC conditions to any of 999 sequences (ABQ65424-ABQ66422) or their
CC fragments. (I) are used to express the corresponding polypeptides (II) or
CC to produce genetically modified plant cells or transgenic plants, which
CC may have improved resistance to disease or stress, or altered
CC metabolic/biosynthetic pathways (for production of commercial,
CC nutritional or medicinal products), or generally any trait of interest,
CC or can be used to screen for biologically active agents (e.g. fungicides,
CC insecticides and antibiotics).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=999909770143.
XX
SQ Sequence 751 BP; 195 A; 164 C; 183 G; 209 T; 0 other;

Query Match 11.4%; Score 89.6; DB 24; Length 751;
Best Local Similarity 56.6%; Pred. No. 2.1e-16;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;

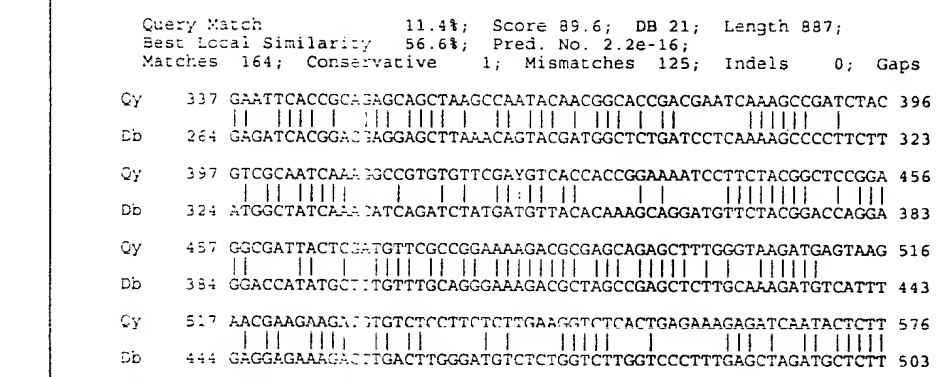
```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          ||  |||| |  ||| |||| |  || ||| | ||| | ||      ||||||  |
Db    186 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAGCCCCTTCTT 245

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | ·| ||||||   |   | |  ||:|| ||     | |     |||||||  | |||
Db    246 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTACGGACCAGGA 305

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||  |  |||| || || |||||||| ||| ||||| | |  ||||||
Db    306 GGACCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAGATGTCATTT 365

Qy    517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTT 576
          | ''|  ||||  || ||     | |    |||||   |       ||| |  || |||||
Db    366 GAGGAGAAAGACTTGACTTGGGATGTCTCTGGTCTTGGTCCCTTTGAGCTAGATGCTCTT 425

Qy    577 AATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT 626
          | |||||||||  ||| ||       ||||| ||   |||||   |||
Db    426 CAAGATTGGGAGTACAAGTTCATGAGCAAGTATGCTAAGGTTGGTACTGT 475
```

RESULT 14
AAC50344
ID AAC50344 standard; DNA; 887 BP.
XX
AC AAC50344;
XX
DT 19-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 64487.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.4%; Score 89.6; DB 21; Length 887;
Best Local Similarity 56.6%; Pred. No. 2.2e-16;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;



RESULT 15
AAC34295
ID AAC34295 standard; DNA; 889 BP.
XX
AC AAC34295;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6138.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX

OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.4%; Score 89.6; DB 21; Length 889;
Best Local Similarity 56.6%; Pred. No. 2.2e-16;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          ||  |||| |  ||| |||| |  || ||| | ||| | ||      ||||||  |
Db    264 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAGCCCCTTCTT 323

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||||   |   | |  ||:|| ||     | |     ||||||||  | |||
Db    324 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTACGGACCAGGA 383

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||  | |||| || || |||||||| ||| ||||| | |  ||||||
Db    384 GGACCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAGATGTCATTT 443

Qy    517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTT 576
          | ||  |||| || ||     | |    |||||   |      ||| |  || |||||
Db    444 GAGGAGAAAGACTTGACTTGGGATATCTCTGGTCTTGGTCCCTTTGAGCTAGATGCTCTT 503

Qy    577 AATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT 626
          | |||||||||  ||| ||       |||||| ||   |||||   |||
Db    504 CAAGATTGGGAGTACAAGTTCATGAGCAAGTATGCTAAGGTTGGTACTGT 553
```

Search completed: January 8, 2003, 14:09:48
Job time : 271 secs

; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 4.2%; Score 33.4; DB 4; Length 1378;
Best Local Similarity 55.7%; Pred. No. 1.3;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 658 TATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTT 717
Db 1355 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1296

Qy 718 TGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAA 772
Db 1295 TCACTCTGATAAATTTATTAAAATAAGAACAAGCATCAGTTGTACATAGAAATAA 1241

RESULT 14
US-08-646-273-18
; Sequence 18, Application US/08646273
; Patent No. 6066502
; GENERAL INFORMATION:
; APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, Thomas, Schmidt,
; APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Heinz.
; TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: WordPerfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,273
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/03706
; FILING DATE: 11-NOV-1994
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA for mRNA
US-08-646-273-18

Query Match 4.2%; Score 33.4; DB 3; Length 1703;
Best Local Similarity 53.6%; Pred. No. 1.5;
Matches 111; Conservative 1; Mismatches 92; Indels 3; Gaps 2;

Qy 360 CCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTT 419
Db 691 CAAAGAATACCTGAGCAAGGTCTCCACCCTCATCAACAGCACAGACAAATGCCTGCTGAA 750

Qy 420 CGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGG 479

Db 751 CAACTACATGATCTGGAA--CCTGGTACGGAAGACGAGCTCCTTCCTCGATCAGCGCTTC 808

Qy 480 AAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTA-AGAACGAAGAAGATGTGTCTCCTT 538
Db 809 CAGGACGCCGACGAGAAGTTCATGGAAGTCATGTATGGGACCAAGAAGACGTGTCTTCCC 868

Qy 539 CTCTTGAAGGTCTCACTGAGAAAGAGA 565
Db 869 CGCTGGAAGTTTTGTGTGAGTGATACA 895

RESULT 15
US-08-646-273-22
; Sequence 22, Application US/08646273
; Patent No. 6066502
; GENERAL INFORMATION:
; APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, Thomas, Schmidt,
; APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Heinz.
; TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: WordPerfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,273
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/03706
; FILING DATE: 11-NOV-1994
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA for mRNA
US-08-646-273-22

Query Match 4.2%; Score 33.4; DB 3; Length 2129;
Best Local Similarity 53.6%; Pred. No. 1.7;
Matches 111; Conservative 1; Mismatches 92; Indels 3; Gaps 2;

Qy 360 CCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTT 419
Db 922 CAAAGAATACCTGAGCAAGGTCTCCACCCTCATCAACAGCACAGACAAATGCCTGCTGAA 981

Qy 420 CGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGG 479
Db 982 CAACTACATGATCTGGAA--CCTGGTACGGAAGACGAGCTCCTTCCTCGATCAGCGCTTC 1039

Qy 480 AAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTA-AGAACGAAGAAGATGTGTCTCCTT 538
Db 1040 CAGGACGCCGACGAGAAGTTCATGGAAGTCATGTATGGGACCAAGAAGACGTGTCTTCCC 1099

Qy 539 CTCTTGAAGGTCTCACTGAGAAAGAGA 565
Db 1100 CGCTGGAAGTTTTGTGTGAGTGATACA 1126

Search completed: January 8, 2003, 14:12:20
Job time : 83 secs

GenCore version 5.1.3

CM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 12:00:08 ; Search time 55 Seconds
(without alignments)
4399.418 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	91.2	11.6	655	3	US-09-188-930-119	Sequence 119, App
	2	62.2	7.9	672	4	US-09-565-808-3	Sequence 3, Appli
	3	55.2	7.0	1868	2	US-08-960-022-5	Sequence 5, Appli
	4	54.4	6.9	588	4	US-09-565-808-1	Sequence 1, Appli
	5	54.4	6.9	788	3	US-08-822-264-2	Sequence 2, Appli
c	6	38.2	4.8	7218	1	US-08-232-463-14	Sequence 14, Appl
c	7	36.2	4.6	4550	4	US-09-462-136-1	Sequence 1, Appli
c	8	35.2	4.5	2323	4	US-09-149-476-24	Sequence 24, Appl
c	9	35	4.4	249	4	US-09-134-001C-1663	Sequence 1663, Ap
c	10	35	4.4	383	4	US-09-221-298-67	Sequence 67, Appl
	11	33.6	4.3	6830	2	US-08-822-445-1	Sequence 1, Appli
	12	33.6	4.3	6830	4	US-09-396-540-1	Sequence 1, Appli
c	13	33.4	4.2	1378	4	US-09-149-476-208	Sequence 208, App
	14	33.4	4.2	1703	3	US-08-646-273-18	Sequence 18, Appl
	15	33.4	4.2	2129	3	US-08-646-273-22	Sequence 22, Appl
	16	33.4	4.2	2314	3	US-08-646-273-29	Sequence 29, Appl
	17	33.4	4.2	2889	1	US-08-289-112-1	Sequence 1, Appli
	18	32.8	4.2	11478	3	US-08-981-803-29	Sequence 29, Appl
	19	32.8	4.2	11478	4	US-08-983-440-29	Sequence 29, Appl
c	20	32.4	4.1	479	4	US-09-030-607-182	Sequence 182, App
c	21	32.4	4.1	479	4	US-09-605-785-182	Sequence 182, App
c	22	32.4	4.1	479	4	US-09-439-313-182	Sequence 182, App
c	23	32.4	4.1	479	4	US-09-352-616A-182	Sequence 182, App
c	24	32.4	4.1	479	4	US-09-232-149A-182	Sequence 182, App
c	25	32.4	4.1	11478	3	US-08-981-803-29	Sequence 29, Appl
c	26	32.4	4.1	11478	4	US-08-983-440-29	Sequence 29, Appl
	27	32	4.1	1413	1	US-08-104-445-1	Sequence 1, Appli
	28	31	4.1	3784	1	US-07-825-033-1	Sequence 1, Appli
c	29	31.8	4.0	1459	2	US-08-970-133-2	Sequence 2, Appli
c	30	31.8	4.0	1459	4	US-09-294-545-2	Sequence 2, Appli
	31	31.8	4.0	8257	4	US-09-484-970B-65	Sequence 65, Appl
c	32	31.6	4.0	152331	3	US-09-128-155-16	Sequence 16, Appl
c	33	31.6	4.0	176373	3	US-09-128-155-17	Sequence 17, Appl
	34	31.6	4.0	1500	4	US-09-593-711A-10	Sequence 10, Appl
	35	31.2	4.0	938	4	US-09-078-294-16	Sequence 16, Appl
c	36	31.2	4.0	3556	1	US-07-971-624E-1	Sequence 1, Appli
c	37	31.2	4.0	3572	1	US-07-971-624E-2	Sequence 2, Appli
	38	31.2	4.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	39	30.8	3.9	610	4	US-08-612-274D-52	Sequence 52, Appl
c	40	30.8	3.9	610	4	US-08-952-014C-52	Sequence 52, Appl
c	41	30.8	3.9	5185	4	US-08-971-395-5	Sequence 5, Appli
c	42	30.8	3.9	162450	4	US-09-345-882-1	Sequence 1, Appli
	43	30.6	3.9	262	4	US-09-437-457-13	Sequence 13, Appl
c	44	30.6	3.9	1944	2	US-08-392-806A-3	Sequence 3, Appli
c	45	30.6	3.9	1944	4	US-09-257-490-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-188-930-119
; Sequence 119, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ontyst, Rene
; APPLICANT: Morrison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11010.1011C1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 655
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-119

Query Match 11.6%; Score 91.2; DB 3; Length 655;
Best Local Similarity 60.7%; Pred. No. 4.8e-19;
Matches 147; Conservative 1; Mismatches 94; Indels 0; Gaps 0;

```
Qy    340 TTCACCGCAGAGTAGCTAAGTCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTC 399
          ||||||| ||| ||||   ||  |||| ||||   || ||      | || ||||||| |
Db    136 TTCACCGAGGAGGAGCTGGCCCGCTACAGCGGCGAGGAGGAGGATCAACCCATCTACTTG 195

Qy    400 GCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGC 459
          ||| | || ||    |||||||:|||||| | || ||    || || ||    |||| |
Db    196 GCAGTGAAGGGAGTGGTGTTCGATGTCACCTCTGGGAAGGAGTTTTATGGACGTGGAGCC 255

Qy    460 GATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAAC 519
            |||      || ||||| || ||| |||||||||| | |||  ||||||     | |
Db    256 CCCTACAACGCCTTGGCCGGGAAGGACTCGAGCAGAGGTGTGGCCAAGATGTCGCTGGAT 315

Qy    520 GAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAAT 579
          | |||  |  ||| |   ||   |||||||||| || ||| |  |  | ||  ||
Db    316 CCTGCAGACCTCACTCATGACATTTCTGGTCTCACTGCCAAGGAGCTGGAAGCCCTCGAT 375

Qy    580 GA 581
          ||
Db    376 GA 377
```

RESULT 2
US-09-565-808-3

```
; Sequence 3, Application US/09565808
; Patent No. 6432674
; GENERAL INFORMATION:
;  APPLICANT:  Hirata, Yuichi
;  TITLE OF INVENTION:  STEROID HORMONE BINDING PROTEIN
;  FILE REFERENCE:  06501-059001
;  CURRENT APPLICATION NUMBER: US/09/565,808
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  WO/JP98/05010
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  JP/9/322376
;  PRIOR FILING DATE:  1997-11-07
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(669)
US-09-565-808-3

  Query Match             7.9%;  Score 62.2;  DB 4;  Length 672;
  Best Local Similarity   54.6%;  Pred. No. 7e-10;
  Matches  171;  Conservative    1;  Mismatches  129;  Indels   12;  Gaps    2;

Qy    320 CTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACG 379
          ||   | ||| |||  ||| |||| |   |||||||  |||| ||| ||||| ||  |
Db    287 CTCGCATGAAGAAGCGGGACTTCAGCTTGGAGCAGCTGCGCCAGTACGACGGCTCCCGCA 346

Qy    380 AATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAAT 439
          |  |   ||  |||    ||||  |||| ||    || |||||:|| ||||  || |
Db    347 ACCC---GCGCATCCTGCTCGCGGTCAATGGGAAAGTCTTCGACGTGACCAAAGGCAGCA 403

Qy    440 CCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTT 499
            |||||||||| | |  ||    ||    || || || || |  || ||   |||||
Db    404 AGTTCTACGGCCCGGCGGGTCCATATGGAATATTTGCTGGTAGGGATGCCTCCAGAGGAC 463

Qy    500 TGGGTAAGATGAGTAAGAACGAAGA---------AGATGTGTCTCCTTCTCTTGAAGGTC 550
          |||  |   |  |     |  ||||         |||||  | |  |  |||   || |
Db    464 TGGCCACATTTTGCCTAGATAAAGATGCACTTAGAGATGAATATGATGATCTCTCAGATT 523

Qy    551 TCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATC 610
          | | ||    | | ||  | | | ||   || ||||| |   | |||  ||  || |||
Db    524 TGAATGCAGTACAAATGGAGAGTGTTCGAGAATGGGAAATGCAGTTTAAAGAAAAATATG 583

Qy    611 CTGTCGTTGGCCG 623
            |   || ||| |
Db    584 ATTATGTAGGCAG 596


RESULT 3
US-08-960-022-5
; Sequence 5, Application US/08960022
; Patent No. 5976837
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John M.
;    APPLICANT:  LaVallie, Edward R.
;    APPLICANT:  Racie, Lisa A.
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    APPLICANT:  Agostino, Michael J.
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/960,022
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sprunger, Suzanne A.
;      REGISTRATION NUMBER:  41,323
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8284
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1868 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-960-022-5

  Query Match             7.0%;  Score 55.2;  DB 2;  Length 1868;
  Best Local Similarity   53.7%;  Pred. No. 2e-07;
  Matches  158;  Conservative    3;  Mismatches  121;  Indels   12;  Gaps    2;

Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || ||||||  | |   ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    265 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 321

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   |||||||:|| ||||  ||        ||||||:|  ||| :
Db    322 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACSGGCCCGAR 381

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |   | | |  |  | |   |
Db    382 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 441

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||          |||| ||    |  |||   |  |||||||      ||
Db    442 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 501

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| ||||||  |  | ||       |||||||  |  ||| |||
Db    502 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 555


RESULT 4
US-09-565-808-1
; Sequence 1, Application US/09565808
; Patent No. 6432674
; GENERAL INFORMATION:
;  APPLICANT:  Hirata, Yuichi
;  TITLE OF INVENTION:  STEROID HORMONE BINDING PROTEIN
;  FILE REFERENCE:  06501-059001
;  CURRENT APPLICATION NUMBER: US/09/565,808
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER:  WO/JP98/05010
;  PRIOR FILING DATE:  1998-11-06
;  PRIOR APPLICATION NUMBER:  JP/9/322376
;  PRIOR FILING DATE:  1997-11-07
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
```

LOCATION: (1)...(585)
S-09-565-808-1

Query Match 6.9%; Score 54.4; DB 4; Length 588;
Best Local Similarity 54.1%; Pred. No. 1.9e-07;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |    ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    214 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 270

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||        ||||||||  |||
Db    271 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 330

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |  | | |   |  |  |  |
Db    331 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTCTGCCTG 390

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||           |||| ||   |  |||  | |||||||      ||
Db    391 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 450

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| ||||||  |  | ||       |||||| |  ||| |||
Db    451 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 504
```

RESULT 5
US-08-822-264-2
; Sequence 2, Application US/08822264
; Patent No. 6033869
; GENERAL INFORMATION:
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL HUMAN CYTOKINE/STEROID
; TITLE OF INVENTION: RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,264
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0233 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CONU.UT101
; CLONE: 2504333
US-08-822-264-2

Query Match 6.9%; Score 54.4; DB 3; Length 788;
Best Local Similarity 54.1%; Pred. No. 2.2e-07;
Matches 159; Conservative 1; Mismatches 122; Indels 12; Gaps 2;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
          || |||||| | |    ||||  | |  | | |||||  | | ||  |   ||  ||   |
Db    313 GACTTCACCCCCGCCGAGCTGCGGCGCTTCGACGGCGTCCAGGACCC---GCGCATACTC 369

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
          | || ||||| |||   ||||||||:|| ||||  ||        ||||||||  |||
Db    370 ATGGCCATCAACGGCAAGGTGTTCGATGTGACCAAAGGCCGCAAATTCTACGGGCCCGAG 429

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
          ||    ||   | | || || |||| ||| ||   ||| |  | | |   |  |  |  |
Db    430 GGGCCGTATGGGGTCTTTGCTGGAAGAGATGCATCCAGGGGCCTTGCCACATTTTGCCTG 489

Qy    517 AACGAAGAA---------GATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATC 567
          |  | |||           |||| ||   |  |||  | |||||||      ||
Db    490 GATAAGGAAGCACTGAAGGATGAGTACGATGACCTTTCTGACCTCACTGCTGCCCAGCAG 549

Qy    568 AATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGC 621
          | ||||| | ||| ||||||  |  | ||       |||||| |  ||| |||
Db    550 GAGACTCTGAGTGACTGGGAGTCTCAGTTCACTTTCAAGTATCATCACGTGGGC 603
```

RESULT 6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 4.8%; Score 38.2; DB 1; Length 7218;
Best Local Similarity 2.1%; Pred. No. 0.1;
Matches 7; Conservative 192; Mismatches 140; Indels 0; Gaps 0;

Qy 265 GGTGAGAGCGAGAAGACGAAGAGAACCTTTTCCCTTGAGAAACAATTGA[illegible]AGAGCTTTA 324
Db 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1336

Qy 325 AAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCA[illegible]GACGAATCA 384
Db 1335 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1276

Qy 385 AAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGG[illegible]AATCCTTC 444
Db 1275 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1216

Qy 445 TACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGT 504
Db 1215 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1156

Qy 505 AAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACT[illegible]AGAAAGAG 564
Db 1155 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1096

Qy 565 ATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCT 603
Db 1095 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCT 1057

RESULT 7
US-09-462-136-1/c
; Sequence 1, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstea, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3837)
US-09-462-136-1

Query Match 4.6%; Score 36.2; DB 4; Length 455[illegible];
Best Local Similarity 48.3%; Pred. No. 0.34;
Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 11 AAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTCTGAAGAAAGA 70
Db 864 AAAAAATGCTCCAAAAAACACAAGCAAAAACGCCATGTAGGTGATCCACATGATGACATA 805

Qy 71 TGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTAGAGGCATTAA 130
Db 804 CATGGCGTCCAAGCCAAGGATCGTCCAGGGAGCAGGAGGAGGTGGGGGCTGGGGCTTGGG 745

Qy 131 AAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAATCAACATCTCC 190

Db 744 GCCACAGACAATAGAGCAGTCTTGGCAGCTACATGGTGCTGTGACCTCATCCACAGACTC 685

Qy 191 GGAACAACGTTAGATCTGTTTCTCAAGGG 219
Db 684 GTCACAGCCTTTGGTGGCATTGTTCATGG 656

RESULT 8
US-09-149-476-24/c
; Sequence 24, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

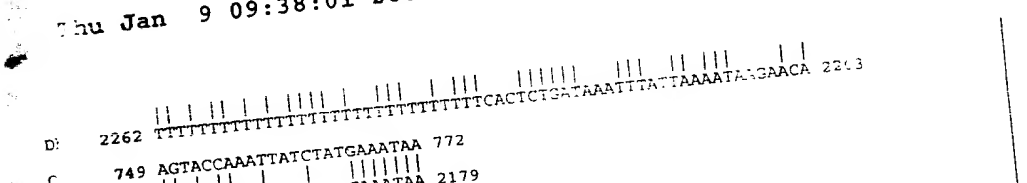
Query Match 4.5%; Score 35.2; DB 4; Length 2323;
Best Local Similarity 52.8%; Pred. No. 0.47;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 629 TCTCTTAGGTCTCTCTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAGGATCT 688

Db 2322 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2263

Qy 689 TTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATC

Thu Jan 9 09:38:01 2003



```
D  2262 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCACTCTGATAAATTTATTAAAATAAGAACA 2203
C   749 AGTACCAAATTATCTATGAAATAA 772
T  2202 AGCATCAGTTGTACATAGAAATAA 2179
```

SULT 9
-09-134-001C-1663/c
Sequence 1663, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING T STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1663
LENGTH: 249
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1663

Query Match 4.4%; Score 35; DB 4; Length 249;
Best Local Similarity 50.9%; Pred. No. 0.15;
Matches 83; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy   624 TGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATGTTATGTAACTATTGTGTGTGAG 683
Db   244 TATCATCATTTGAATTTTTATTTTGTTGTTGTTCTTTGTCGGTATAAT ATATTGT GT 185
Qy   684 GATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGATCTGATCGTTTTGATACAATTA 743
Db   184 GATCTACATGATTCGAATTTTGTTCTTTTGAATTTTTCTTTCTACGTTTTTCTTCTAATA 125
Qy   744 CCATAAGTACCAAATTATCTATGAAATAAATCGGGGATTTCGT 786
Db   124 CGCTTAGTACAATAAACAAAATAATTAATATTGGAGATAACAT 82
```

RESULT 10
US-09-221-298-67/c
; Sequence 67, Application US/09221298
; Patent No. 6284241
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER
; FILE REFERENCE: 210121.471
; CURRENT APPLICATION NUMBER: US/09/221,298
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (32)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (46)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (59)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (132)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (190)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (234)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)
; OTHER INFORMATION: Where n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (272)
; OTHER INFORMATION: Where n is a, c, g or t
US-09-221-298-67

Query Match 4.4%; Score 35; DB 4; Length 383;
Best Local Similarity 50.0%; Pred. No. 0.2;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
Qy   431 CCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGA 490
Db   383 CCGAAAAGGAGTTTTATGGACGAGGAGCCCCTTACAATGCCTTGACGGGGAAGGACTCCA 324
Qy   491 GCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTC 550
Db   323 CTAGAGGGGTAGCCAAGATGTCCTTGGATCCTGCAGACCTCACCCATGACANTACGGGTN 264
Qy   551 TCACTGAGAAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATT 596
Db   263 TCACGGCCAAGGAACTGGAGGCCCTGGATNAGGTTTTCACCAAAGT 218
```

RESULT 11
US-08-822-445-1
; Sequence 1, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
;

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 8, 2003, 14:09:55 ; Search time 2802 Seconds
(without alignments)
8194.901 Million cell updates/sec

Title: US-09-649-866A-1
Perfect score: 789
Sequence: 1 atcatcaacaaaaacaattc.....taaatcggggatttcgtgtt 789

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	489.2	62.0	522	8	AY084294	AY084294 Arabidops
c 2	489.2	62.0	103495	8	AC006585	AC006585 Arabidops
3	487.6	61.8	2836	8	AY090994	AY090994 Arabidops
4	282.8	35.8	477	8	AY085799	AY085799 Arabidops
c 5	282.8	35.8	83922	8	T9A4	AF096373 Arabidops
6	282.8	35.8	99856	8	ATF24G24	AL049488 Arabidops
7	282.8	35.8	199861	8	ATCHRIV29	AL161517 Arabidops
c 8	125.2	15.9	139336	2	AP005115	AP005115 Oryza sat
c 9	122	15.5	133158	2	AP004054	AP004054 Oryza sat
10	113.8	14.4	83922	8	T9A4	AF096373 Arabidops
c 11	113.8	14.4	99856	8	ATF24G24	AL049488 Arabidops
c 12	113.8	14.4	199861	8	ATCHRIV29	AL161517 Arabidops
13	90.2	11.4	141382	2	AC130811	AC130811 Medicago
14	89.6	11.4	847	8	AF153284	AF153284 Arabidops
15	89.6	11.4	889	8	AY086811	AY086811 Arabidops
c 16	84.6	10.7	81609	8	AC027035	AC027035 Arabidops
c 17	84.6	10.7	97242	8	AC051630	AC051630 Arabidops
18	79.2	10.0	861	8	AY046006	AY046006 Arabidops
19	75.8	9.6	702	6	AX412363	AX412363 Sequence
20	75.8	9.6	702	8	AY079104	AY079104 Arabidops
21	75.8	9.6	724	9	AF173937	AF173937 Homo sapi
22	75.8	9.6	930	8	AF153283	AF153283 Arabidops
23	75.8	9.6	945	8	AF419567	AF419567 Arabidops
24	75.8	9.6	953	9	BC008823	BC008823 Homo sapi
25	74.6	9.5	697	9	AK074431	AK074431 Homo sapi
c 26	70.6	8.9	110000	2	LMFLCHR36_23	Continuation (24 c
27	70	8.9	504	9	AF049672	AF049672 Homo sapi
28	69.2	8.8	306	8	LEPT52	X59882 L.esculentu
29	65.8	8.3	509	8	MSA248337	AJ248337 Medicago
30	65	8.2	48420	8	AB025603	AB025603 Arabidops
31	62.2	7.9	1081	6	AX454748	AX454748 Sequence
32	62.2	7.9	1081	6	AX491226	AX491226 Sequence
33	62.2	7.9	1874	9	HSAJ2030	AJ002030 Homo sapi
34	62.2	7.9	1897	9	BC016692	BC016692 Homo sapi
35	61.6	7.8	919	8	AY084353	AY084353 Arabidops
36	59.2	7.5	63554	8	H0410G08	AL512546 Oryza sat
c 37	59.2	7.5	73187	2	OSIG00033	AL732338 Oryza sat
c 38	59.2	7.5	159525	8	OSJN00019	AL606444 Oryza sat
39	57.4	7.3	2708	9	AK091741	AK091741 Homo sapi
40	56.4	7.1	1490	3	AY061163	AY061163 Drosophil
41	56	7.1	1893	4	SSSTERMBP	X99714 S.scrofa mR
42	55.2	7.0	1868	6	AR083260	AR083260 Sequence
43	54.4	6.9	1866	9	BC034238	BC034238 Homo sapi
44	54.4	6.9	1890	6	AX335970	AX335970 Sequence
45	54.4	6.9	1890	6	AX336196	AX336196 Sequence

ALIGNMENTS

RESULT 1
AY084294
LOCUS AY084294 522 bp mRNA linear PLN 21-JUN-2002
DEFINITION Arabidopsis thaliana clone 10261 mRNA, complete sequence.
ACCESSION AY084294
VERSION AY084294.1 GI:21403004
KEYWORDS FLI_CDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 522)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

```
  TITL      Full-length messenger RNA sequences greatly improve genome
            annotation
  JOUR  L   Genome Biol. (2002) In press
REFERE  E   2  (bases 1 to 522)
  AUTH  S   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITL      Full-Length cDNA from Arabidopsis thaliana
  JOUR  L   Unpublished
REFERE  E   3  (bases 1 to 522)
  AUTH  S   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITL      Direct Submission
  JOUR  L   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road
            Malibu, CA 90265, USA
  COMM      This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATU  S              Location/Qualifiers
     ce  irce         1. .522
                      /organism="Arabidopsis thaliana"
                      /db_xref="taxon:3702"
                      /clone="10261"
         S            54. .356
                      /codon_start=1
                      /product="putative steroid binding protein"
                      /protein_id="AAM60885.1"
                      /db_xref="GI:21536553"
                      /translation="MEFTAEQLSQYNGTDESKPIYVAIKGRVFDVTTGKSFYGSGPDY
                      SMFAGKDASRALGKMSKNEEDVSPSLEGLTEKEINTLNDWETKFEAKYPVVGRVVS"
ASE   UNT      141 a    100 c    121 g    160 t
RIGI
```

```
 Que  Match            62.0%;  Score 489.2;  DB 8;  Length 522;
 Best Local Similarity  99.0%;  Pred. No. 6e-120;
 Mat  es  491;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      94 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 73

Qy      54 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCCCAATCAAAGGCCG 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      74 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 133

Qy     414 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
           |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 193

Qy     474 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     194 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 253

Qy     534 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     254 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 313

Qy     594 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     314 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 373
```

```
Qy     654 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     374 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 433

Qy     714 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     434 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 493

Qy     774 TCGGGGATTTCGTGTT 789
           ||||||||||||||||
Db     494 TCGGGGATTTCGTGTT 509
```

```
RESULT 2
AC006585/c
LOCUS       AC006585               103495 bp    DNA     linear   PLN 11-MAR-2002
DEFINITION  Arabidopsis thaliana chromosome 2 clone F27C12 map mi238, complete
            sequence.
ACCESSION   AC006585
VERSION     AC006585.8  GI:20197873
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 103495)
  AUTHORS   Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,
            Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
            Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
            Nierman,W.C., Fraser,C.M. and Venter,J.C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 103495)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 103495)
  AUTHORS   Town,C.D. and Kaul,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT     On Apr 18, 2002 this sequence version replaced gi:6598610.
            Sequencing, analysis, and annotation were performed within the
             CSHL/WUGSC/ABI Arabidopsis Genome Sequencing Consortium.
             Information on physical mapping and YAC and BAC library
             construction as well as added annotation can be viewed at
             http://www.cshl.org/arabweb/. We used GenScan, Grail, and MZEF for
             predicting coding exons and assembling genes. BAC F6P23 maps to
            YAC
             yUPCIC8B1.
FEATURES             Location/Qualifiers
     source          1. .103495
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /map="mi238"
                     /clone="F27C12"
     misc_feature    complement(1. .10309)
                     /note="overlap with BAC clone F13D4 (AL031369:1. .10306)."
     repeat_region   complement(100. .136)
                     /rpt_family="AT_rich"
     repeat_region   complement(206. .236)
                     /rpt_family="AT_rich"
     gene            1028. .3697
                     /gene="At2g25070"
                     /note="F27C12.1; supported by full length cDNA:
                     Ceres:11428"
     mRNA            join(1028. .1197,1624. .1742,1826. .1889,2061. .2169,
                     2257. .2410,2506. .2755,2850. .2926,3010. .3062,3149. .3203,
                     3287. .3697)
                     /gene="At2g25070"
```

```
CDS             join(1626. .1742,1826. .1889,2061. .2169,2257. .2410,
                2506. .2755,2850. .2926,3010. .3062,3149. .3203,3287. .3475)
                /gene="At2g25070"
                /codon_start=1
                /product="putative protein phosphatase 2C"
                /protein_id="AAD23006.1"
                /db_xref="GI:4559345"
                /translation="MGTYLSSPKTEKLSEDGENDKLRFGLSSMQGWRATMEDAHAAIL
                DLDDKTSFFGVYDGHGGKVVAKFCAKYLHQQVISNEAYKTGDVETSLRRAFFRMDDMM
                QGQRGWRELAVLGDKMNKFSGMIEGFIWSPRSGDTNNQPDSWPLEDGPHSDFTGPTSG
                CTACVALIKDKKLFVANAGDSRCVISRKSQAYNLSKDHKPDLEVEKERILKAGGFIHA
                GRINGSLNLTRAIGDMEFKQNKFLPSEKQMVTADPDINTIDLCDDDDFLVVACDGIWD
                CMSSQELVDFIHEQLKSETKLSTVCEKVVDRCLAPDTATGEGCDNMTIILVQFKKPNP
                SETEPEDSKPEPSEDEPSSSS"
gene            complement(4062. .4666)
                /gene="At2g25060"
                /note="F27C12.2"
mRNA            complement(join(<4062. .4414,4489. .>4666))
                /gene="At2g25060"
CDS             complement(join(4062. .4414,4489. .4666))
                /gene="At2g25060"
                /codon_start=1
                /product="similar to early nodulins"
                /protein_id="AAD23007.1"
                /db_xref="GI:4559346"
                /translation="MASSSLHVAIFSLIFLFSLAAANEVTVGGKSGDWKIPPSSSYSF
                TEWAQKARFKVGDFIVFRYESGKDSVLEVTKEAYNSCNTTNPLANYTDGETKVKLDRS
                GPFYFISGANGHCEKGQKLSLVVISPRHSVISPAPSPVEFEDGPALAPAPISGSVRLG
                GCYVVLGLVLGLCAWF"
gene            7609. .10233
                /gene="At2g25050"
                /note="F27C12.3"
mRNA            join(<7609. .8283,8382. .8603,8902. .>10233)
                /gene="At2g25050"
CDS             join(7609. .8283,8382. .8603,8902. .10233)
                /gene="At2g25050"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAD23008.1"
                /db_xref="GI:4559347"
                /translation="MLEDEDYRVYVSRIMSQLREQFPGASFMVFNFRDGDSRSRMESV
                LTEYDMTIMDYPRHYEGCPLLTMETVHHFLKSAESWLLLSQQNILLSHCELGGWPTLA
                FMLASLLLYRKQFSGEHRTLEMIYKQAPRELLQLMSPLNPLPSQLRFLQYISRRNVGS
                QWPPLDQALTLDCVNLRLIPDFDGEGGCRPIFRIYGQDPFMASDRTSKVLFSMPKRSK
                AVRQYKQADCELVKIDINCHILGDVVLECITLGSDLEREEMMFRVVFNTAFLRSNILT
                LNRGEIDVLWNTTDRFPKDFSAEVIFSEMGAGKKLASVDLPHMEEKDVLPMEAFAKVQ
                EIFSEAEWLDPNSDVAVTVFNQITAANILQESLDSGSPRSPDSRSLLESALEKVKEKT
                KLMISENIVSSPDTSSPEKEKDTMSSHKSYADPNSILKKVDESRGLRVSVQRNVHSKI
                FSPRMVQSPVTSPLPNRSPTQGSPASISRFHSSPSSLGITSILHDHGSCKDEESTSSS
                PASPSISFLPTLHPLTSSQPKKASPQCPQSPTPVHSNGPPSAEAAVTSSPLPPLKPLR
                ILSRPPPPPPPPPISSLRSTPSPSSTSNSIATQGPPPPPPPPPLQSHRSALSSSPLPP
                PLPPKKLLATTNPPPPPPPPLHSNSRMGAPTSSLVLKSPPVPPPPAPAPLSRSHNGNI
                PPVPGPPLGLKGRGILQNLKGQGQTRKANLKPYHWLKLTRAVQGSLWAEAQKSDEAAT
                YD"
repeat_region   complement(9805. .9887)
                /rpt_family="(GGA)n"
repeat_region   complement(14951. .14971)
                /rpt_family="AT_rich"
gene            complement(15196. .18244)
                /gene="At2g25030"
                /note="F27C12.4; similar to HSP100 ClpB GI:9651530
                (Phaseolus lunatus)"
mRNA            complement(join(<15196. .15393,16835. .17074,17335. .17525,
                17815. .17997,18097. .>18244))
                /gene="At2g25030"
CDS             complement(join(15196. .15393,16835. .17074,17335. .17525,
                17815. .17997,18097. .18244))
                /gene="At2g25030"
                /codon_start=1
                /product="HSP100/ClpB, putative"
                /protein_id="AAD23010.2"
                /db_xref="GI:20197876"
                /translation="MLEQILHERIIAQDLDVESVADAIRCSKAGISDPNRLIASFMFM
                IPSVVSQLVGASPGYVGYGDGGKLTEVVRRRPYSVVQFDEIEKPHLDVFSILLQLLD
                GRITNSHGSLHILETIRNNEDIKEAFYEMMKQQVVELARKTFRPEFMNRIDEYIVSQ
                LNSSEISKIVELQVVQMRQVKKRLEQNKINLEYTKEAVDLLAQLGFDPNNGARPVKQ
                VTEKLVKKEITLKVLKGDFAEDGTILIDADQPNNKLTDAKHEALETYGSDLTKMARQG
                KLPPLIGRDDEVNRCIQILCRMTESNPVIIGEPGVGKTAIVEG"
repeat_region   complement(16306. .16351)
                /rpt_family="AT_rich"
repeat_region   complement(18672. .18728)
                /rpt_family="AT_rich"
repeat_region   complement(18996. .19052)
                /rpt_family="AT_rich"
repeat_region   complement(20060. .20210)
                /rpt_family="(TAAAAA)n"
repeat_region   complement(20211. .20238)
                /rpt_family="AT_rich"
repeat_region   20766. .20786
                /rpt_family="AT_rich"
repeat_region   21200. .21251
                /rpt_family="AT_rich"
repeat_region   complement(21883. .21957)
                /rpt_family="AT_rich"
repeat_region   22547. .22581
                /rpt_family="AT_rich"
repeat_region   22623. .22643
                /rpt_family="AT_rich"
repeat_region   23323. .23358
                /rpt_family="AT_rich"
repeat_region   24133. .24191
                /rpt_family="AT_rich"
gene            complement(25342. .27903)
                /gene="At2g25020"
                /note="F27C12.6; predicted by genefinder"
mRNA            complement(join(<25342. .26334,27106. .27237,27349. .27620,
                27870. .>27903))
                /gene="At2g25020"
CDS             complement(join(25342. .26334,27106. .27237,27349. .27620,
                27870. .27903))
                /gene="At2g25020"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAD23011.1"
                /db_xref="GI:4559350"
                /translation="MDAKRKLSQEKDMYQNDDKRKKVVALENQEAGGSCIDGNVAHWP
                KTKFNGLVNVPIVYLALEENIVKNDLTLRSSLKHVIFGGSSGGSNSGKGLAKDLYLD
                NELVEVHVRSDREINMMKDKGKAIFDLNEAVWPRDEAVSNTFEDHPLPPGPGSNICK
                IKMRVPKDGFSGTGIFPHPLHKMIMKPHHFEIFNFLCKNLVVENSNGCIIAQTPLSE
                TFLMISFIYGYLEKHPNSKPLFVLPKWVLNFWKTKFGELKVNDLVLLDFYSAKASTR
                QQLEVLNRWIKTRSIIFLGAKQFSNIISDNSGAEASDSCRDILLNIPSVVVFDRGTD
                RNEMMSFLKVVARIKTPHKVLLTGSLYQNNIKEVFNILDVVFPEFLKHNRIGKNIRK
                LNVEADGPSTNLKMPLFDKLEEALLSQDSDHGDKISYLTELRMLTNKFIYNHKEEFL
                EVPGLMDFTVVLKPTISQKSAWEIE"
repeat_region   complement(26429. .26467)
                /rpt_family="AT_rich"
repeat_region   complement(26578. .26629)
                /rpt_family="AT_rich"
repeat_region   complement(26763. .26795)
                /rpt_family="AT_rich"
repeat_region   complement(28025. .28050)
                /rpt_family="(TA)n"
repeat_region   complement(28292. .28324)
                /rpt_family="(TA)n"
repeat_region   28601. .28629
                /rpt_family="AT_rich"
repeat_region   complement(29181. .29220)
                /rpt_family="AT_rich"
repeat_region   complement(29476. .29504)
                /rpt_family="AT_rich"
repeat_region   30114. .30157
                /rpt_family="AT_rich"
```

Query Match 62.0%; Score 489.2; DB 8; Length 103495;
Best Local Similarity 99.0%; Pred. No. 1.2e-119;
Matches 491; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy      4 TTCCCTTGAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 353
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]6 TTCTTTCCAGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCA 57526

Qy      4 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 GCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCG 57466

Qy      4 TGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 473
          |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 TGTGTTCGATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTT 57406

Qy      4 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 CGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTC 57346

Qy      4 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 TCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAA 57286

Qy      4 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 ATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATT 57226

Qy      4 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 GCACTATGTTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCG 57166

Qy      4 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  57[illegible]5 TGTTTGGATCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAA 57106

Qy      4 TCGGGGATTTCGTGTT 789
          ||||||||||||||||
Db  57[illegible]5 TCGGGGATTTCGTGTT 57090
```

```
[illegible]SULT
[illegible]909
LOCUS       AY090994               2836 bp    mRNA    linear   PLN 21-APR-2002
DEFINITION  Arabidopsis thaliana At2g24940 mRNA sequence.
ACCESSION   AY090994
VERSION     AY090994.1  GI:20258816
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2836)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2836)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
```

```
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R  and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .2836
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="RAFL09-54-M24 (R19002)"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert.
                     ecotype: Columbia"
     gene            1. .2836
                     /gene="At2g24940"
     misc_feature    1. .2820
                     /gene="At2g24940"
                     /note="This is a potentially unprocessed transcript
                     corresponding to At2g24940.  The cDNA contains the full
                     annotated ORF from At2g24940 (GI:15224648) with no
                     mismatches.  However, the gene was most likely
                     misannotated since this would give the cDNA a 2483 bp
                     3'UTR.  In addition, there is a larger ORF located in
                     different frame starting at position 2170 that may also be
                     utilized."
                     /evidence=experimental
     misc_difference 634
                     /gene="At2g24940"
                     /note="compared to genomic sequence"
                     /replace="c"
     misc_difference 2819
                     /gene="At2g24940"
                     /note="compared to genomic sequence"
                     /replace="t"
BASE COUNT    [illegible]684 a    595 c    596 g    961 t
ORIGIN
```

```
  Query Match             61.8%;  Score 487.6;  DB 8;  Length 2836;
  Best Local Similarity   99.8%;  Pred. No. 2e-119;
  Matches  487;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    302 AGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCC 361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 AGAAACAATTGATCAGAGCTTTAAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCC 62

Qy    362 AATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCG 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 AATACAACGGCACCGACGAATCAAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCG 122

Qy    422 AYGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAA 481
          |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 ATGTCACCACCGGAAAATCCTTCTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAA 182

Qy    482 AAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTC 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTC 242

Qy    542 TTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAG 601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    243 TTGAAGGTCTCACTGAGAAAGAGATCAATACTCTTAATGATTGGGAGACCAAATTTGAAG 302

Qy    602 CTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATG 661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 CTAAGTATCCTGTCGTTGGCCGTGTTGTCTCTTAGGTCTCTCTTCTGAGATTGCACTATG 362

Qy    662 TTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGA 721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 TTATGTAACTATTGTGTGTGAGGATCTTTGTGTTGTGTGTTTTCTGATTTCGTGTTTGGA 422

Qy    722 TCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGAT 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 TCTGATCGTTTTGATACAATTACCATAAGTACCAAATTATCTATGAAATAAATCGGGGAT 482

Qy    782 TTCGTGTT 789
          ||||||||
Db    483 TTCGTGTT 490
```

```
RESULT 4
AY085799
LOCUS       AY085799                 477 bp    mRNA    linear   PLN 25-JUN-2002
DEFINITION  Arabidopsis thaliana clone 1831 mRNA, complete sequence.
ACCESSION   AY085799
VERSION     AY085799.1  GI:21404509
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
  JOURNAL   Genome Biol. (2002) In press
REFERENCE   2  (bases 1 to 477)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-Length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 477)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="1831"
     CDS             72. .344
                     /codon_start=1
                     /product="probable wound-induced protein"
                     /protein_id="AAM63015.1"
                     /db_xref="GI:21553934"
                     /translation="MSSTSKAWTVAVSIGAVEALKDQLGLCRWNYILRSVNQHLRNN.
                     RSVSQGKRFSSSSVSAAVTSSGESEKAKKAEESLRTVMYLSCWGPN"
BASE COUNT      164 a     98 c     86 g    129 t
ORIGIN
```

```
  Query Match           35.8%;  Score 282.8;  DB 8;  Length 477;
  Best Local Similarity  99.3%;  Pred. No. 9.6e-65;
  Matches  284;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0

Qy      1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 62

Qy     61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 122

Qy    121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 182

Qy    181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 242

Qy    241 GTCTCCGCAGTCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGA 286
          |||||||||| |||||||||||||||||||||||||| | ||||
Db    243 GTCTCCGCAGTCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 288
```

```
RESULT 5
T9A4/c
LOCUS       T9A4                     83922 bp    DNA     linear   PLN 06-OCT-1998
DEFINITION  Arabidopsis thaliana BAC T9A4.
ACCESSION   AF096373
VERSION     AF096373.1  GI:3695400
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 83922)
  AUTHORS   Washington University Genome Sequencing Center.
  TITLE     The A. thaliana Genome Sequencing Project
  JOURNAL   Unpublished (1997)
REFERENCE   2  (bases 1 to 83922)
  AUTHORS   Zidanic,M., McQuerry,Y. and Smith,A.
  TITLE     The sequence of A. thaliana T9A4
  JOURNAL   Unpublished (1998)
REFERENCE   3  (bases 1 to 83922)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     Submitted by:
                    Genome Sequencing Center
                    Department of Genetics, Washington University,
                    St. Louis, MO 63108, USA
                    e-mail: rwilson@watson.wustl.edu

            MAPPING: Clones were assigned to the YAC map by hybridization by
            M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted
            by M. Marra, WashU, to pick the best candidates for sequencing.

            NOTICE: This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.
```

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 3' clone is T17F16. Actual start of this clone is at base position 1 of T9A4 ; actual end is at 83992 of T9A4.

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder(P. Green and L. Hillier, ms in preparation).

```
FEATURE             Location/Qualifiers
     source         1. .83922
                    /organism="Arabidopsis thaliana"
                    /cultivar="Columbia"
                    /db_xref="taxon:3702"
                    /chromosome="IV"
                    /map="unknown"
                    /clone="T9A4"
     gene           486. .3161
                    /gene="T9A4.4"
     CDS            join(486. .665,756. .839,938. .1147,1257. .1382,1469. .1606,
                    1698. .1801,1893. .1944,2103. .2191,2298. .2481,2533. .2596,
                    2672. .2739,2908. .3030,3126. .3161)
                    /gene="T9A4.4"
                    /note="similar to isoleucyl-tRNA synthetases"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62806.1"
                    /db_xref="GI:3695406"
                    /translation="MVLSTALFEKPAFKNLICNGLVLAEDGKKMAKKLRNYPPPLEVI
                    DEYGAVSFPYNMLLCCDAVRLYLINSPVVRAEPLRFKKEGVLGVVKDVFLPWYNAYRF
                    LVQNAKRLETEGGVPFVPTDLATIQSANILDQWIHSATQSLVRFVREEMDAYRLYTVV
                    PRLLKFLDNLTNIYVRFNRKRLKGRTGEDDCHTALSTLFNVLLTSCKVMAPFTPFFTE
                    TLYQNLRKACKGSEESVHYCSIPPREGMEGERIELSVTRMMKIIDLARNIRERNKLPL
                    KTPLKEMIVVHPDADFLNDITGYVLEELNVRSLVPCNDTLKYASLKAEPDFSVLGKRL
                    GKSMGLVAKEVKEMSQKDILAFEEAGEVTIANHLLKETDIKVSHAMMFSCEIALRIVR
                    VFKRPDDLKENEIDSAGDGDVLVILDLRADDSLVEAGFAREIVNRIQKLRKKSGLEPT
                    DFVEVYFQSLDEDESVSKQVLVSQLWIKRKFMFPT"
     gene           complement(5712. .9371)
                    /gene="T9A4.5"
     CDS            complement(join(5712. .5939,6786. .6918,8299. .8988,
                    9046. .9371))
                    /gene="T9A4.5"
                    /note="similar to potassium transport proteins"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62807.1"
                    /db_xref="GI:3695407"
                    /translation="MDRVVAKIAKIRSQLTKLRSLFFLYFIYFLFFSFLGFLALKITK
                    PRTTSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIFLTILMFLGGEIFTSFLNL
                    YVSYFTNSIEDRCDVETVTDYREGLIKIDERASKCLYSVVLSYHLVTNLVGSVLLLVY
                    VNFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTNENMIIFRKNSGLIWLLIPQ
                    VLMGNTLFPCFLVLLIWGLYKITKRDEYGYILKCHNKMGYSHLLSVRLCVLLGVTVLG
                    FLIIQLLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSRHTGETIVDLSTLSPAILVL
                    FILMIGLIVSQLSFLTICIFLISITERQNLQRDPINFNVLNITLEVIRYFCNGSAYGN
                    VGFTTGYSCERRVDISDGGCKDASYGFAGRWSPMGKFVLIIVMFYGRFKQFTAKSGRA
                    WILYPSSS"
     gene           complement(26302. .26574)
                    /gene="T9A4.6"
     CDS            complement(26302. .26574)
                    /gene="T9A4.6"
                    /note="contains similarity to Solanum lycopersicum
                    (tomato) wound-induced protein (GB:X59882)"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62808.1"
                    /db_xref="GI:3695408"
                    /translation="MSSTSKAWTVAVSIGAVEALKDQLGLCRWNYILRSVNQHLRNNV
                    RSVSQGKRFSSSSVSAAVTSSGESEKAKKAEESLRTVMYLSCWGPN"
     gene           28609. .29893
                    /gene="T9A4.3"
     CDS            join(28609. .28785,28860. .29270,29345. .29479,29558. .29674,
                    29759. .29893)
                    /gene="T9A4.3"
                    /note="contains similarity to the pfkB family of
                    carbohydrate kinases (Pfam: PF00294, E=1.6e-75)"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62803.1"
                    /db_xref="GI:3695403"
                    /translation="MANTPLIVSFGEMLIDFVPDTSGVSLAESTGFLKAPGGAPANVA
                    CAITKLGGKSAFIGKFGDDEFGHMLVNILKKNGVNSEGVCFDTNARTALAFVTLKKDG
                    EREFMFYRNPSADMLLKESELNKDLIKKAKIFHYGSISLISEPCRTAHMAAMKTAKDA
                    GVLLSYDPNVRLPLWPSTEAAIEGIKSIWNEADIIKVSDDEVTFLTRGDAEKDDVVLS
                    LMHDKLKLLIVTDGEKGCRYYTKKFKGRVPGYAVKAVDTTGAGDSFVGAFLVSLGKDG
                    SILDDEGKLKEALAFANACGAVCTTQKGAIPALPTPADAQKLMKSKSK"
     gene           complement(30255. .30842)
                    /gene="T9A4.7"
     CDS            complement(30255. .30842)
                    /gene="T9A4.7"
                    /note="contains similarity to heat shock hsp20 proteins
                    (Pfam: PF00011, E=1.2e-46"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62802.1"
                    /db_xref="GI:3695402"
                    /translation="MMKHLLSIFFIGALLLGNIKTSEGSLSSALETTPGSLLSDLWLD
                    RFPDPFKILERIPLGLERDTSVALSPARVDWKETAEGHEIMLDIPGLKKDEVKIEVEE
                    NGVLRVSGERKREEEKKGDQWHRVERSYGKFWRQFKLPDNVDMESVKAKLENGVLTIN
                    LTKLSPEKVKGPRVVNIAAEEDQTAKISSSESKEL"
     gene           31853. .32412
                    /gene="T9A4.2"
     CDS            join(31853. .32053,32125. .32412)
                    /gene="T9A4.2"
                    /note="contains similarity to Arabidopsis thaliana
                    salt-tolerance protein (GB:X95572) and CONSTANS-like 1
                    proteins"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62805.1"
                    /db_xref="GI:3695405"
                    /translation="MKIQCEVCEKAEAEVLCCSDEAVLCKPCDIKVHEANKLFQRHHR
                    VALQKDAASATTASGAPLCDICQERKGYFFCLEDRAMLCNDCDEAIHTCNSHQRFLLS
                    GVQVSDQSLTENSECSTSFSSETYQIQSKVSLNSQYSSEETEAGNSGEIVHKNPSVIL
                    SP"
     gene           complement(34194. .35118)
                    /gene="T9A4.8"
     CDS            complement(join(34194. .34922,35026. .35118))
                    /gene="T9A4.8"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62809.1"
                    /db_xref="GI:3695409"
                    /translation="MTSILSAVSKLRGCVNQIENKNPSGASEEDILNQAKMLLTQYEK
                    YKRGFKFDHVWPILKGIEKFANDNMKTPPAFQGEGRDVTSSSSFSINTESSPSPGMNS
                    IDLNMDSEDANFSLSSRPMGLKKAKRKQQSEEQFKQLLEQNDKLIKAITKGTSERNEI
                    QRQKIEVARMKEENKILFADLNSISDPSSRAYVENERKRILEKRAQTNQHEEDGEGSQ
                    YHGSQYRASHYQESLFHGKQVQGEPDQGEDKRSPNNQEDFTQYYNYLSGTGNNFP"
     gene           complement(36300. .39184)
                    /gene="T9A4.9"
     CDS            complement(join(36300. .36656,37387. .37539,38026. .38214,
                    38746. .38838,38999. .39184))
                    /gene="T9A4.9"
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC62813.1"
                    /db_xref="GI:3695413"
                    /translation="MKLGCALPMTCMILACYIICGSLLMSHCHGVIKEAKTLKSNEDL
                    EIEHKLKLINKPAFKIVKTINGERYGCVDFYKQPGLDHSSMKNHTFHHKTHRNTTFGH
```

```
                     FWENGVGCPIGTVPIPRVTKDALLRMKSFDSDNSNPQSSWSKTYKPASSIDDHHFAVV
                     RTTKGTRSYNGASMNINTFTPSVGPMQFSASRMHFQIGNEFIQVGWIDKINGNWWLLM
                     GTSWEEVGFWPSSRFKESSGTMVEWGGEVYSPSPPNPPMGNSHYPKGSPKVDSYVRLI
                     TTVDENYNTDKTVKNTERYSDSCYKVRDATETFWSHVGHLIIYGGPGCK"
     gene            complement(41254. .44127)
                     /gene="T9A4.10"
     CDS             complement(join(41254. .41451,42287. .42442,42949. .43137,
                     43675. .43767,43984. .44127))
                     /gene="T9A4.10"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAC62810.1"
                     /db_xref="GI:3695410"
                     /translation="MCMILFCYILCCSLLMSHCHGVVEAAKALKSNEDLEIEQKLELI

  Query Match             35.8%;  Score 282.8;  DB 8;  Length 83922;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-64;
  Matches  284;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26643 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 26584

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26583 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 26524

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26523 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 26464

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   26463 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 26404

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGA 286
          |||||||||||||||||||||||||||||||||||| | ||||
Db   26403 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 26358


RESULT 6
ATF24G24
LOCUS       ATF24G24               99856 bp    DNA     linear   PLN 27-AUG-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, BAC clone  F24G24 (ESSA
            project).
ACCESSION   AL049488
VERSION     AL049488.1  GI:4538949
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 99856)
  AUTHORS   Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
            Mayer,K.F.X. and Schueller,C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 99856)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosome 4 can be
            viewed at: http://websvr.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .99856
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     misc_feature    1. .4596
                     /note="overlap to BAC F28M11; please refer to this entry
                     for analysis and annotation"
     gene            2780. .5818
                     /gene="F24G24.10"
     CDS             join(2780. .2971,3173. .3265,3758. .3991,4498. .4657,
                     4909. .5027,5489. .5818)
                     /gene="F24G24.10"
                     /note="similarity to IB1C3-1 protein, Arabidopsis
                     thaliana, AJ011845"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB39774.1"
                     /db_xref="GI:4538950"
                     /translation="MKMKLGYSLPVMCMILFCYILCCSLLMSHCHGVVEAAKALKSNE
                     DLEIEQKLELINKHTVKIIKCTNGERYGCVDFYKQPGLDHSLMKNHTFHHKMRLMSYP
                     EGSKIKKQTHINKTFGHFWKNGVGRPIGTVPILLVSKEALLKMKSFDGDNSNPQSSWS
                     KTYKPTSSNGGHHFAVVRTTKGKPRRYNGVAMNINSFNPPVGPMEFSAGRMHFQIGNE
                     FVQVGWTVRGHGCYNPLCPVGSGIILVSHEVTPGLLTKHNDPELSIIKDKIYGHWWLL
                     MGNSSSSTWKEIGFWPTHRFKESFGTGVEWGGEVYSPASTSPPMGNSHFPKGSPKIDS
                     YVRLITTWDENYGLDMVVKNTERFSNSCYKVKDAQSRFGQM"
     exon            2780. .2971
                     /gene="F24G24.10"
                     /number=1
     intron          2972. .3172
                     /gene="F24G24.10"
                     /number=1
     exon            3173. .3265
                     /gene="F24G24.10"
                     /number=2
     intron          3266. .3757
                     /gene="F24G24.10"
                     /number=2
     exon            3758. .3991
                     /gene="F24G24.10"
                     /number=3
     intron          3992. .4497
                     /gene="F24G24.10"
                     /number=3
     exon            4498. .4657
                     /gene="F24G24.10"
                     /number=4
     intron          4658. .4908
                     /gene="F24G24.10"
                     /number=4
     exon            4909. .5027
                     /gene="F24G24.10"
                     /number=5
     intron          5028. .5488
                     /gene="F24G24.10"
                     /number=5
     exon            5489. .5818
                     /gene="F24G24.10"
                     /number=6
     gene            8685. .10641
                     /gene="F24G24.20"
     CDS             join(8685. .8915,9402. .9554,10285. .10641)
                     /gene="F24G24.20"
                     /note="similarity to IB1C3-1 protein, Arabidopsis
                     thaliana, AJ011845"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB39775.1"
                     /db_xref="GI:4538951"
                     /translation="MRMSYSEGSKMKRKTHRNTTFGHFWENGVGCPIGTVPIPRVTKD
                     ALLRMKSFDSDNSNPQSSWSKTYKPASSIDDHHFAVVRTTKGTRSYNGASMNINTFTP
                     SVGPMQFSASRMHFQIGNEFIQVGWIDKINGNWWLLMGTSWEEVGFWPSSRFKESSGT
                     MVEWGGEVYSPSPPNPPMGNSHYPKGSPKVDSYVRLITTVDENYNTDKTVKNTERYSD
                     SCYKVRDATETFWSHVGHLIIYGGPGCK"
     exon            8685. .8915
                     /gene="F24G24.20"
```

```
                /number=1
intron          8916. .9401
                /gene="F24G24.20"
                /number=1
exon            9402. .9554
                /gene="F24G24.20"
                /number=2
intron          9555. .10284
                /gene="F24G24.20"
                /number=2
exon            10285. .10641
                /gene="F24G24.20"
                /number=3
gene            11823. .12747
                /gene="F24G24.30"
CDS             join(11823. .11915,12019. .12747)
                /gene="F24G24.30"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAB39776.1"
                /db_xref="GI:4538952"
                /translation="MTSILSAVSKLRGCVNQIENKNPSGASEEDILNQAKMLLTQYEK
                YKRGFKFDHVWPILKGIEKFANDNMKTPPAFQGEGRDVTSSSSFSINTESSPSPGMNS
                IDLNMDSEDANFSLSSRPMGLKKAKRKQQSEEQFKQLLEQNDKLIKAITKGTSERNEI
                QRQKIEVARMKEENKILFADLNSISDPSSRAYVENERKRILEKRAQTNQHEEDGEGSQ
                YHGSQYRASHYQESLFHGKQVQGEPDQGEDKRSPNNQEDFTQYYNYLSGTGNNFP"
exon            11823. .11915
                /gene="F24G24.30"
                /number=1
intron          11916. .12018
                /gene="F24G24.30"
                /number=1
exon            12019. .12747
                /gene="F24G24.30"
                /number=2
gene            14529. .15088
                /gene="F24G24.40"
gene            complement(join(14529. .14816,14888. .15088))
                /gene="F24G24.40"
CDS             complement(join(14529. .14816,14888. .15088))
                /gene="F24G24.40"
                /note="strong similarity to zinc-finger protein R2931,
                Oryza sativa, PIR3:JE0116"
                /codon_start=1
                /product="zinc-finger-like protein"
                /protein_id="CAB39777.1"
                /db_xref="GI:4538953"
                /translation="MKIQCEVCEKAEAEVLCCSDEAVLCKPCDIKVHEANKLFQRHHR
                VALQKDAASATTASGAPLCDICQERKGYFFCLEDRAMLCNDCDEAIHTCNSHQRFLLS
                GVQVSDQSLTENSECSTSFSSETYQIQSKVSLNSQYSSEETEAGNSGEIVHKNPSVIL
                SP"
exon            complement(14529. .14816)
                /gene="F24G24.40"
                /number=1
intron          complement(14817. .14887)
                /gene="F24G24.40"
                /number=1
exon            complement(14888. .15088)
                /gene="F24G24.40"
                /number=2
gene            16099. .16686
                /gene="F24G24.50"
CDS             16099. .16686
                /gene="F24G24.50"
                /note="contains EST gb:AA597970"
                /codon_start=1
                /product="heat shock protein 22.0"
                /protein_id="CAB39778.1"
                /db_xref="GI:4538954"
                /translation="MMKHLLSIFFIGALLLGNIKTSEGSLSSALETTPGSLLSDLWLD
                RFPDPFKILERIPLGLERDTSVALSPARVDWKETAEGHEIMLDIPGLKKDEVKIEVEE
                NGVLRVSGERKREEEKKGDQWHRVERSYGKFWRQFKLPDNVDMESVKAKLENGVLTIN
                LTKLSPEKVKGPRVVNIAAEEDQTAKISSSESKEL"
exon            16099. .16686
                /gene="F24G24.50"
                /number=1
gene            17048. .18332
                /gene="F24G24.60"
gene            complement(join(17048. .17182,17267. .17383,17462. .17596,
                17671. .18081,18156. .18332))
                /gene="F24G24.60"
CDS             complement(join(17048. .17182,17267. .17383,17462. .17596,
                17671. .18081,18156. .18332))
                /gene="F24G24.60"
                /EC_number="2.7.1.4"
                /note="strong similarity to fructokinase, Lycopersicon
                esculentum, gb:U62329
                Contains pfkB family of carbohydrate kinases signatures,
                Pfkb_Kinases_1 [GGAPANVACAITKLGGKSAFIGKFG], Pfkb_Kinases_2
                [DTTGAGDSFVGAFL]"
                /codon_start=1
                /product="fructokinase-like protein"
                /protein_id="CAB39779.1"
                /db_xref="GI:4538955"
                /translation="MANTPLIVSFGEMLIDFVPDTSGVSLAESTGFLKAPGGAPANVA
                CAITKLGGKSAFIGKFGDDEFGHMLVNILKKNGVNSEGVCFDTNARTALAFVTLKKDG
```

```
  Query Match          35.8%;  Score 282.8;  DB 8;  Length 99856;
  Best Local Similarity  99.3%;  Pred. No. 1.9e-64;
  Matches  284;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20298 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 20357

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20358 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 20417

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20418 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 20477

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   20478 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 20537

Qy     241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGACGAAGA 286
          |||||||||||||||||||||||||||||||||||| | ||||
Db   20538 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 20583


RESULT 7
ATCHRIV29
LOCUS       ATCHRIV29              199861 bp    DNA     linear   PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 29.
ACCESSION   AL161517
VERSION     AL161517.2  GI:7267723
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 155039)
  AUTHORS   Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 147497 to 199861)
  AUTHORS   Wedler,H., Wedler,E., Wambutt,R., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 199861)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
```

```
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV28 at the 5' end and an
            overlap with ATCHRIV30 at the 3' end.
FEATURES            Location/Qualifiers
     source         1. .199861
                    /organism="Arabidopsis thaliana"
                    /variety="Columbia"
                    /db_xref="taxon:3702"
                    /chromosome="4"
     gene           7035. .7307
                    /gene="AT4g10270"
     CDS            7035. .7307
                    /gene="AT4g10270"
                    /note="strong similarity to wound-induced protein,
                    Lycopersicon esculentum, PIR2:S19773
                    contains EST gb:AI995575.1"
                    /codon_start=1
                    /product="probable wound-induced protein"
                    /protein_id="CAB78150.1"
                    /db_xref="GI:7267724"
                    /translation="MSSTSKAWTVAVSIGAVEALKDQLGLCRWNYILRSVNQHLRNIV
                    RSVSQGKRFSSSSVSAAVTSSGESEKAKKAEESLRTVMYLSCWGPN"
     exon           7035. .7307
                    /gene="AT4g10270"
                    /number=1
     gene           7450. .8226
                    /gene="AT4g10280"
     gene           complement(join(7450. .7666,8021. .8226))
                    /gene="AT4g10280"
     CDS            complement(join(7450. .7666,8021. .8226))
                    /gene="AT4g10280"
                    /note="similarity to predicted protein, Arabidopsis
                    thaliana"
                    /codon_start=1
                    /product="putative protein"
                    /protein_id="CAB78151.1"
                    /db_xref="GI:7267725"
                    /translation="MGNHFLTLSLLLVTVCVCVSFITTKLNPKEAIVSVSSDSEIPTE
                    IHGVKILRQASDTKLAQLGVASWPKWEGAPSKFPWEFKKTETIYFMEGKVKVNVDGYD
                    EEEETFEIGKGDVVVFPKDMKVVWEITEAVKKQYSLEE"
     exon           complement(7450. .7666)
                    /gene="AT4g10280"
                    /number=1
     intron         complement(7667. .8020)
                    /gene="AT4g10280"
                    /number=1
     exon           complement(8021. .8226)
                    /gene="AT4g10280"
                    /number=2
     tRNA           10468. .10619
                    /product="tRNA-Ser"
                    /note="codon recognized: GGA"
     gene           12188. .12418
                    /gene="AT4g10290"
     gene           complement(12188. .12418)
                    /gene="AT4g10290"
     CDS            complement(12188. .12418)
                    /gene="AT4g10290"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="CAB78152.1"
                    /db_xref="GI:7267726"
                    /translation="MPRKFPWKFKKTETMYFVEGKLKVKVEDHHKEGEALEFVAGDLV
                    VFPQDMNVFVDVIEDVKKRYYRESEIEESELP"
     exon           complement(12188. .12418)
                    /gene="AT4g10290"
                    /number=1
     gene           16794. .18176
                    /gene="AT4g10300"
     CDS            join(16794. .16993,17972. .18176)
                    /gene="AT4g10300"
                    /note="similarity to predicted protein, Arabidopsis
                    thaliana
                    contains EST gb:AI823169.1"
                    /codon_start=1
                    /product="putative protein"
                    /protein_id="CAB78153.1"
                    /db_xref="GI:7267727"
                    /translation="MGCIGVVNVTSINPYLTRRSNKPYNSRRPSSMAAAIRAESTEKL
                    SITIEKNPPESKLTQLGVRSWPKWGCPPSKFPWTYSAKETCYLLQGKVKVYPNGSDEG
                    VEIEAGDFVVFPKGMSCTWDVSVAVDKHYQPE"
     exon           16794. .16993
                    /gene="AT4g10300"
                    /number=1
     intron         16994. .17971
                    /gene="AT4g10300"
                    /number=1
     exon           17972. .18176
                    /gene="AT4g10300"
                    /number=2
     gene           24238. .27814
                    /gene="AT4g10310"
     CDS            join(24238. .25310,26586. .26819,27652. .27814)
                    /gene="AT4g10310"
                    /note="similarity to potassium transport protein TRK1,
                    Saccharomyces cerevisiae, PIR2:JU0466"
                    /codon_start=1
                    /product="potassium uptake transporter-like protein"
                    /protein_id="CAB78154.1"
                    /db_xref="GI:7267728"
                    /translation="MDRVVAKIAKIRSQLTKLRSLFFLYFIYFLFFSFLGFLALKITK
                    PRTTSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIFLTILMFLGGEIFTSFLNL
                    VVSYFTKFVFPHNKIRHILGSYNSDSSIEDRCDVETVTDYREGLIKIDERASKCLYSV
                    VLSYHLVTNLVGSVLLLVYVNFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTN
                    ENMIIFRKNSGLIWLLIPQVLMGNTLFPCFLVLLIWGLYKITKRDEYGYILKNHNKMG
                    YSHLLSVRLCVLLGVTVLGFLIIQLLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSR
                    HTGETIVDLSTLSPAILVLFILMMYLPPYTLFMPLTEQKTIEKEGGDDDSENGKKVKK
                    SGLIVSQLSFLTICIFLISITERQNLQRDPINFNVLNITLEVIRIINKISFVMAVHME
                    TLVSLPGTAVNGVWTSAMVAAKTRVMGLQDDGVQWENSY"
     exon           24238. .25310
                    /gene="AT4g10310"
                    /number=1
     intron         25311. .26585
                    /gene="AT4g10310"
                    /number=1
     exon           26586. .26819
                    /gene="AT4g10310"
                    /number=2
     intron         26820. .27651
                    /gene="AT4g10310"
                    /number=2
     exon           27652. .27814
                    /gene="AT4g10310"
                    /number=3
     gene           29756. .36739
                    /gene="AT4g10320"
     gene           complement(join(29756. .30244,30328. .30393,30579. .30701,
                    30870. .30937,31013. .31076,31128. .31311,31418. .31506,
                    31653. .31716,31808. .31911,32003. .32140,32227. .32352,
                    32462. .32671,32770. .32853,32977. .33136,33651. .33784,
                    33863. .33952,34221. .34379,34479. .34557,34914. .35011,
                    35156. .35269,35441. .35962,36051. .36116,36206. .36739,)
                    /gene="AT4g10320"
     CDS            complement(join(29756. .30244,30328. .30393,30579. .30701,
                    30870. .30937,31013. .31076,31128. .31311,31418. .31506,
                    31653. .31716,31808. .31911,32003. .32140,32227. .32352,
                    32462. .32671,32770. .32853,32977. .33136,33651. .33784,
                    33863. .33952,34221. .34379,34479. .34557,34914. .35011,
                    35156. .35269,35441. .35962,36051. .36116,36206. .36739,)
```

```
                     /gene="AT4g10320"
                     /note="strong similarity to isoleucine--tRNA ligase, Homo
                     sapiens, PIR2:I59314
                     Contains Aminoacyl-transfer RNA synthetases class-I
                     signature AA49-60;Prokaryotic membrane lipoprotein lipid
                     attachment site AA788-798;Prokaryotic membrane lipoprotein
                     lipid attachment site AA1215-1225"
                     /codon_start=1
                     /product="isoleucine-tRNA ligase-like protein"
                     /protein_id="CAB78155.1"
                     /db_xref="GI:7267729"
                     /translation="MEEVCEGKEFSFPRQEEDVLSFWTEIDAFKTQLKRTENLPEYIF
                     YDGPPFATGLPHYGHILAGTIKDIVTRYQTMTGHHVTRRFGWDCHGLPVENEIDRKLN
                     IKRRDEVIKMGIDKYNEECRSIVTRYVAEWEKVITRCGRWIDFKNDYKTMDLPFMESV
                     WWVFSQLWEKNLVYRGFKVMPYSTGCKTPLSNFEAGQNYKEVEDPEIMVAFPVIGQD
                     NAAFVAWTTTPWTLPSNLALCVNAKFVYVKVRNKNTGKVYIVAESRLSALFTDKPKAK
                     LSNGPAGDTKKANPKAKGAKPESAADSYEVLEKFNGASLVGKKYEPLFDYFSDFSSEA
                     FRVVADDYVTDDSGTGIVHCAPAFGEDDYRVCLLNKIIKKGENLVVAVDDDGLFTERI
                     THFSGRYVKDADKDIIEAVKAKGRLVKTGSFTHSYPFCWRSDTPLIYRAVPSWFVRVE
                     QLKEKLLKSNEQTEWVPGYVKDKRFHNWLENARDWAISRSRFWGTPLPIWISDDGEEV
                     VIMDSVEKLEKLSGVKVFDLHRHHIDHITIPSSRGDEFGVLRRVEDVFDCWFESGSMP
                     YAYIHYPFENKELFEKNFPGDFVAEGLDQTRGWFYTLMVLSTALFEKPAFKNLICNGL
                     VLAEDGKKMAKKLRNYPPPLEVIDEYGADAVRLYLINSPVVRAEPLRFKKEGVLGVVK
                     DVFLPWYNAYRFLVQNAKRLETEGGVPFVPTDLATIQSANILDQWIHSATQSLVREFVR
                     EEMDAYRLYTVVPRLLKFLDNLTNIYVRFNRKRLKGRTGEDDCHTALSTLFNVLLTSC
                     KVMAPFTPFFTETLYQNLRKACKGSEESVHYCSIPPREGMEGERIELSVTRMMKIIDL
                     ARNIRERNKLPLKTPLKEMIVVHPDADFLNDITGVLREYVLEELNVRSLVPCNDTLKY
                     ASLKAEPDFSVLGKRLGKSMGLVAKEVKEMSQKDILAFEEAGEVTIANHLLKETDIKV
                     SHAMMFSCEIALRIVRVFKRPDDLKENEIDSAGDGDVLVILDLRADDSLVEAGFAREI
                     VNRIQKLRKKSGLEPTDFVEVYFQSLDEDESVSKQVLVSQEQIKDSIGSTLLLSSLM
                     PSHAVIIADETFTPKETSDESVKKVPKLSYKISLARPALKFNEEAVLALYSGESFSHH
                     NRFSLSILSRFLFNVYIYGSRRCEICNRASNILVIKGSFEFEVRVPSRRWEGSLVKLL
```

```
Query Match             35.8%;  Score 282.8;  DB 8;  Length 139861;
Best Local Similarity   99.3%;  Pred. No. 2.1e-64;
Matches  284;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

      1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   6966 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 7025

     61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   7026 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 7085

    121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   7086 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 7145

    181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   7146 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 7205

    241 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 286
        |||||||||||||||||||||||||||||||||||||| | ||||
   7206 GTCTCCGCAGCCGTTACCTCCTCTGGTGAGAGCGAGAAGGCTAAGA 7251


RESULT 8
AP005115/c
LOCUS       AP005115               139336 bp    DNA     linear   HTG 20-APR-2002
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 2 clone P0700F06,
            *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AP005115
VERSION     AP005115.1  GI:20219003
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
            clone:P0700F06.
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE     Cryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
            clone:P0700F06
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 139336)
  AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .139336
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="P0700F06"
BASE COUNT    40464 a  28725 c  29241 g  40806 t    100 others
ORIGIN

  Query Match             15.9%;  Score 125.2;  DB 2;  Length 139336;
  Best Local Similarity   63.3%;  Pred. No. 2.8e-22;
  Matches  180;  Conservative    1;  Mismatches  109;  Indels    0;  Gaps    0;

Qy     324 AAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATC 383
           | |||     |  |||  | || || | ||||||  |    ||| |||| | ||||  |||
Db   62759 AGAGATGGCGGCGGAGCTGACGGCGGCGCAGCTGCGGGCGTACGACGGGAGCGACCCATC 62700

Qy     384 AAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTT 443
           ||||||||||||||| |  ||   |||   || | |||:|||||| ||||      |||
Db   62699 GAAGCCGATCTACGTCTCCGTCCGCGGCAAGGTCTACGACGTCACCTCCGGCCGCGGCTT 62640

Qy     444 CTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGG 503
           ||||||| |||| ||||  ||| |  | ||||||||    || || ||| | ||  | ||
Db   62639 CTACGGCCCCGGCGGCGCCTACGCCGTCTTCGCCGGCCGCGAGGCCAGCCGCGCGCTCGG 62580

Qy     504 TAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGA 563
           ||||||   ||| |||| |  || || ||       ||    || ||| | || || ||
Db   62579 CAAGATGTCCAAGGACGACGCCGACGTCTCCGGCGACCTCTCCGGCCTCTCCGACAAGGA 62520

Qy     564 GATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCG 623
           | ||      ||    || |||||||||| ||  | || ||||| || ||||| | |||
Db   62519 GCTCGGCGTCCTCGCCGACTGGGAGACCAAGTTCCAGGCCAAGTACCCCGTCGTCGCCCG 62460


RESULT 9
AP004054/c
LOCUS       AP004054               133158 bp    DNA     linear   HTG 21-MAR-2002
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 2 clone
            OJ1249_F12, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AP004054
VERSION     AP004054.1  GI:15208422
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
            clone:OJ1249_F12.
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
```

```
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:OJ1249_F12
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 133158)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) Takuji Sasaki, National Institute of
            Agrobiological Resources, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .133158
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="2"
                     /clone="OJ1249_F12"
BASE COUNT    38255 a  27854 c  27875 g  39121 t     53 others
ORIGIN

  Query Match             15.5%;  Score 122;  DB 2;  Length 133158;
  Best Local Similarity   63.3%;  Pred. No. 2e-21;
  Matches  190;  Conservative    1;  Mismatches  109;  Indels    0;  Gaps    0;

Qy     324 AAAGAAAAAGATGGAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATC 383
           | |||     |  |||  | || || | ||||||  |     ||| |||| | ||||  |||
Db   19147 AGAGATGGCGGCGGAGCTGACGGCGGCGCAGCTGCGGGCGTACGACGGGAGCGACCCATC 19088

Qy     384 AAAGCCGATCTACGTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGA-AATCCTT 442
           ||||||||||||||  |  ||   |||   || | |||:|||||| ||||      |||
Db   19087 GAAGCCGATCTACGTCTCCGTCCGCGGCAAGGTCTACGACGTCACCTCCGGCGCGGCTT 19028

Qy     444 CTACGGCTCCGGAGGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGG 503
           |||||||  |||| ||||  |||  |  | ||||||||     || || ||| | ||  | ||
Db   19027 CTACGGCCCCGGCGGCGCCTACGCCGTCTTCGCCGGCCGCGAGGCCAGCCGCGCGCTCGG 18968

Qy     504 TAAGATGAGTAAGAACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGA 563
           ||||||   ||| |||| |  || || ||        ||    || ||| | || || ||
Db   18967 CAAGATGTCCAAGGACGACGCCGACGTCTCCGGCGACCTCTCCGGCCTCTCCGACAAGGA 18908

Qy     564 GATCAATACTCTTAATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCG 623
           | ||      ||    || |||||||||| ||  | || ||||| || ||||| | |||
Db   18907 GCTCGGCGTCCTCGCCGACTGGGAGACCAAGTTCCAGGCCAAGTACCCCGTCGTCGCCCG 18848


RESULT 10
T9A4
LOCUS       T9A4                    83922 bp    DNA     linear   PLN 06-OCT-1998
DEFINITION  Arabidopsis thaliana BAC T9A4.
ACCESSION   AF096373
VERSION     AF096373.1  GI:3695400
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 83922)
  AUTHORS   Washington University Genome Sequencing Center.
  TITLE     The A. thaliana Genome Sequencing Project
  JOURNAL   Unpublished (1997)
REFERENCE   2  (bases 1 to 83922)
  AUTHORS   Didanic,M., McQuerry,Y. and Smith,A.
  TITLE     The sequence of A. thaliana T9A4
  JOURNAL   Unpublished (1998)
REFERENCE   3  (bases 1 to 83922)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     Submitted by:
                    Genome Sequencing Center
                    Department of Genetics, Washington University,
                    St. Louis, MO 63108, USA
                    e-mail: rwilson@watson.wustl.edu


            MAPPING: Clones were assigned to the YAC map by hybridization by
            M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted
            by M. Marra, WashU, to pick the best candidates for sequencing.


            NOTICE:  This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone


                        NEIGHBORING COSMID INFORMATION:


            The 3' clone is T17F16.  Actual start of this clone is at base
            position 1 of T9A4 ; actual end is at 83992 of T9A4.

            NOTES:

            Coding sequences below are predicted from computer analysis, using
            the program Genefinder(P. Green and L. Hillier, ms in preparation).
FEATURES             Location/Qualifiers
     source          1. .83922
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="unknown"
                     /clone="T9A4"
     gene            486. .3161
                     /gene="T9A4.4"
     CDS             join(486. .665,756. .839,938. .1147,1257. .1382,1469. .1606,
                     1698. .1801,1893. .1944,2103. .2191,2298. .2481,2533. .2596,
                     2672. .2739,2908. .3030,3126. .3161)
                     /gene="T9A4.4"
                     /note="similar to isoleucyl-tRNA synthetases"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAC62806.1"
                     /db_xref="GI:3695406"
                     /translation="MVLSTALFEKPAFKNLICNGLVLAEDGKKMAKKLRNYPPPLEVI
                     EEYGAVSFPYNMLLCCDAVRLYLINSPVVRAEPLRFKKEGVLGVVKDVFLPWYNAYRF
                     LVQNAKRLETEGGVPFVPTDLATIQSANILDQWIHSATQSLVRFVREEMDAYRLYTVV
                     PRLLKFLDNLTNIYVRFNRKRLKGRTGEDDCHTALSTLFNVLLTSCKVMAPFTPFFTE
                     TLYQNLRKACKGSEESVHYCSIPPREGMEGERIELSVTRMMKIIDLARNIRERNKLPL
                     PTPLKEMIVVHPDADFLNDITGYVLEELNVRSLVPCNDTLKYASLPAEPDFSVLGKRL
                     GSMGLVAKEVKEMSQKDILAFEEAGEVTIANHLLKETDIKVSHAAMPSCEIALRIVR
                     PKRPDDLKENEIDSAGDGDVLVILDLRADDSLVEAGFAREIVNRIQKLRKKSGLEPT
                     PVEVYFQSLDEDESVSKQVLVSQLWIKRKFMFPT"
     gene            complement(5712. .9371)
```

```
                /gene="T9A4.5"
CDS             complement(join(5712. .5939,6786. .6918,8299. .8988,
                9046. .9371))
                /gene="T9A4.5"
                /note="similar to potassium transport proteins"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62807.1"
                /db_xref="GI:3695407"
                /translation="MDRVVAKIAKIRSQLTKLRSLFFLYFIYFLFFSFLGFLALKITK
                PRTTSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIFLTILMFLGGEIFTSFLNL
                YVSYFTNSIEDRCDVETVTDYREGLIKIDERASKCLYSVVLSYHLVTNLVGSVLLLVY
                VNFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTNENMIIFRKNSGLIWLLIFQ
                VLMGNTLFPCFLVLLIWGLYKITKRDEYGYILKNHNKMGYSHLLSVRLCVLLGVTVLG
                FLIIQLLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSRHTGETIVDLSTLSPAILVL
                FILMIGLIVSQLSFLTICIFLISITERQNLQRDPINFNVLNITLEVIRYFCNGSAYGN
                VGFTTGYSCERRVDISDGGCKDASYGFAGRWSPMGKFVLIIVMFYGRFKQFTAKSGRA
                WILYPSSS"
gene            complement(26302. .26574)
                /gene="T9A4.6"
CDS             complement(26302. .26574)
                /gene="T9A4.6"
                /note="contains similarity to Solanum lycopersicum
                (tomato) wound-induced protein (GB:X59882)"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62808.1"
                /db_xref="GI:3695408"
                /translation="MSSTSKAWTVAVSIGAVEALKDQLGLCRWNYILRSVNQHLRNNV
                RSVSQGKRFSSSSVSAAVTSSGESEKAKKAEESLRTVMYLSCWGPN"
gene            28609. .29893
                /gene="T9A4.3"
CDS             join(28609. .28785,28860. .29270,29345. .29479,29556. .29674,
                29759. .29893)
                /gene="T9A4.3"
                /note="contains similarity to the pfkB family of
                carbohydrate kinases (Pfam: PF00294, E=1.6e-75)"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62803.1"
                /db_xref="GI:3695403"
                /translation="MANTPLIVSFGEMLIDFVPDTSGVSLAESTGFLKAPNGAPANVA
                CAITKLGGKSAFIGKFGDDEFGHMLVNILKKNGVNSEGVCFDTNARTALAFVTLKKDG
                EREFMFYRNPSADMLLKESELNKDLIKKAKIFHYGSISLISEPCRTAHMAAMKTAKDA
                GVLLSYDPNVRLPLWPSTEAAIEGIKSIWNEADIIKVSDDEVTFLTRGDAEKDDVVLS
                LMHDKLKLLIVTDGEKGCRYYTKKFKGRVPGYAVKAVDTTGAGDSFVGAFLVSLGKDG
                SILDDEGKLKEALAFANACGAVCTTQKGAIPALPTPADAQKLMKSKSK"
gene            complement(30255. .30842)
                /gene="T9A4.7"
CDS             complement(30255. .30842)
                /gene="T9A4.7"
                /note="contains similarity to heat shock hsp20 proteins
                (Pfam: PF00011, E=1.2e-46"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62802.1"
                /db_xref="GI:3695402"
                /translation="MMKHLLSIFFIGALLLGNIKTSEGSLSSALETTPGSLLSDLWLD
                RFPDPFKILERIPLGLERDTSVALSPARVDWKETAEGHEIMLDIPGLKKDEVKIEVEE
                NGVLRVSGERKREEEKKGDQWHRVERSYGKFWRQFKLPDNVDMESVKAKLENGVLTIN
                LTKLSPEKVKGPRVVNIAAEEDQTAKISSSESKEL"
gene            31853. .32412
                /gene="T9A4.2"
CDS             join(31853. .32053,32125. .32412)
                /gene="T9A4.2"
                /note="contains similarity to Arabidopsis thaliana
                salt-tolerance protein (GB:X95572) and CONSTANS-like 1
                proteins"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62805.1"
                /db_xref="GI:3695405"
                /translation="MKIQCEVCEKAEAEVLCCSDEAVLCKFCDIKVHFANKLFQPHHR
                VALQKDAASATTASGAPLCDICQERKGYFFCLEDRAMLCNDCDEAIHTCNSHQRFLLS
                GVQVSDQSLTENSECSTSFSSETYQIQSKVSLNSQYSGEETEAGNSGEIVHKNPSVIL
                SP"
gene            complement(34194. .35118)
                /gene="T9A4.8"
CDS             complement(join(34194. .34922,35026. .35118))
                /gene="T9A4.8"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62809.1"
                /db_xref="GI:3695409"
                /translation="MTSILSAVSKLRGCVNQIENKNPSGASEEDILNQAKMLLTQYEK
                YKRGFKFDHVWPILKGIEKFANDNMKTPPAFQGEGRDVTSSSSFSINTESSPSPGMNS
                IDLNMDSEDANFSLSSRPMGLKKAKRKQQSEEQFKQLLEQNDKLIKAITKGTSERNEI
                QRQKIEVARMKEENKILFADLNSISDPSSRAYVENERKRILEKRAQTNQHEEDGEGSQ
                YHGSQYRASHYQESLFHGKQVQGEPDQGEDKRSPNNQEDFTQYYNYLSGTGNNFP"
gene            complement(36300. .39184)
                /gene="T9A4.9"
CDS             complement(join(36300. .36656,37387. .37539,38026. .38214,
                38746. .38838,38999. .39184))
                /gene="T9A4.9"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62813.1"
                /db_xref="GI:3695413"
                /translation="MKLGCALPMTCMILACYIICGSLLMSHCHGVIKEAKTLKSNEDL
                EIEHKLKLINKPAFKIVKTINGERYGCVDFYKQPGLDHSSMKNHTFHHKTHRNTTFGH
                FWENGVGCPIGTVPIPRVTKDALLRMKSFDSDNSNPQSSWSKTYKPASSIDDHHFAVV
                RTTKGTRSYNGASMNINTFTPSVGPMQFSASRMHFQIGNEFIQVGWIDKINGNWWLLM
                GTSWEEVGFWPSSRFKESSGTMVEWGGEVYSPSPPNPPMGNSHYPKGSPKVDSYVRLI
                TTVDENYNTDKTVKNTERYSDSCYKVRDATETFWSHVGHLIIYGGPGCK"
gene            complement(41254. .44127)
                /gene="T9A4.10"
CDS             complement(join(41254. .41451,42287. .42442,42949. .43137,
                43675. .43767,43984. .44127))
                /gene="T9A4.10"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC62810.1"
                /db_xref="GI:3695410"
                /translation="MCMILFCYILCCSLLMSHCHGVVEAAKALKSNEDLEIEQKLELI
```

```
  Query Match             14.4%;  Score 113.8;  DB 8;  Length 83922;
  Best Local Similarity   70.4%;  Pred. No. 2.9e-19;
  Matches  171;  Conservative    0;  Mismatches   62;  Indels   10;  Gaps    1;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           |||||||||||| ||||     ||  || |  ||  |||  | ||||||
Db   27843 ATCATCAACAAACACAAAATCGAAACCAAATCACTTAAACAAGAAAGAA----------A 27892

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
             |||| | ||||| ||| |||||||| | ||||  || || |  ||||||||||||||
Db   27893 CAAAGAGAAATGAGCTCTGCAAGCAAAACGTGGATGGTTGCTGCAAGCATCGGAGCCGTT 27952

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           |||||||| |||||||||||||  | ||||| |||||||| |  |||| || |  |||
Db   27953 GAGGCATTGAAAGACCAACTAGGCGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAAT 28012

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||  |||| || |||||| | ||||| || || ||||  || | | |||||||| || ||
Db   28013 CAGTATCTACGCAACAACTTAAGATCCGTGTCGCAAGCTAAGAAGCTCTCTTCCTCATCA 28072

Qy     241 GTC 243
           |||
Db   28073 ATC 28075


RESULT 11
ATF24G24/c
LOCUS       ATF24G24               99856 bp    DNA     linear   PLN 27-AUG-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, BAC clone  F24G24 (ESSA
            project).
ACCESSION   AL049488
```

```
VERSION     AL049488.1  GI:4538949
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 99856)
  AUTHORS   Bevan,M., Murphy,G., Ridley,P., Hudson,S., Bancroft,I., Mewes,H.W.,
            Mayer,K.F.X. and Schueller,C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 99856)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosome 4 can be
            viewed at: http://websvr.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .99856
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     misc_feature    1. .4596
                     /note="overlap to BAC F28M11; please refer to this entry
                     for analysis and annotation"
     gene            2780. .5818
                     /gene="F24G24.10"
     CDS             join(2780. .2971,3173. .3265,3758. .3991,4498. .4657,
                     4909. .5027,5489. .5818)
                     /gene="F24G24.10"
                     /note="similarity to IB1C3-1 protein, Arabidopsis
                     thaliana, AJ011845"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB39774.1"
                     /db_xref="GI:4538950"
                     /translation="MKMKLGYSLPVMCMILFCYILCCSLLMSHCHGVVEAAKALKSNE
                     DLEIEQKLELINKHTVKIIKCTNGERYGCVDFYKQPGLDHSLMKNHTFHHKMRLMSYP
                     EGSKIKKQTHINKTFGHFWKNGVGRPIGTVPILLVSKEALLKMKSFDGDNSNPQSSWS
                     KTYKPTSSNGGHHFAVVRTTKGKPRRYNGVAMNINSFNPPVGPMEFSAGRMHFQIGNE
                     FVQVGWTVRGHGCYNPLCPVGSGIILVSHEVTPGLLTKHNDFELSIIKDKIYGHWWLL
                     MGNSSSSTWKEIGFWPTHRFKESFGTGVEWGGEVYSPASTSPPMGNSHFPKGSPKIDS
                     YVRLITTWDENYGLDMVVKNTERFSNSCYKVKDAQSRFGQM"
     exon            2780. .2971
                     /gene="F24G24.10"
                     /number=1
     intron          2972. .3172
                     /gene="F24G24.10"
                     /number=1
     exon            3173. .3265
                     /gene="F24G24.10"
                     /number=2
     intron          3266. .3757
                     /gene="F24G24.10"
                     /number=2
     exon            3758. .3991
                     /gene="F24G24.10"
                     /number=3
     intron          3992. .4497
                     /gene="F24G24.10"
                     /number=3
     exon            4498. .4657
                     /gene="F24G24.10"
                     /number=4
     intron          4658. .4908
                     /gene="F24G24.10"
                     /number=4
     exon            4909. .5027
                     /gene="F24G24.10"
                     /number=5
     intron          5028. .5488
                     /gene="F24G24.10"
                     /number=5
     exon            5489. .5818
                     /gene="F24G24.10"
                     /number=6
     gene            8685. .10641
                     /gene="F24G24.20"
     CDS             join(8685. .8915,9402. .9554,10285. .10641)
                     /gene="F24G24.20"
                     /note="similarity to IB1C3-1 protein, Arabidopsis
                     thaliana, AJ011845"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB39775.1"
                     /db_xref="GI:4538951"
                     /translation="MRMSYSEGSKMKRKTHRNTTFGHFWENGVGCPIGTVPIPRVTKD
                     ALLRMKSFDSDNSNPQSSWSKTYKPASSIDDHHFAVVRTTKGTRSYNGASMNINSFTP
                     SVGPMQFSASRMHFQIGNEFIQVGWIDKINGNWWLLMGTSWEEVGFWPSSRFKESSGT
                     MVEWGGEVYSPSPPNPPMGNSHYPKGSPKVDSYVRLITTVDENYNTDKTVKNTERYSD
                     SCYKVRDATETFWSHVGHLIIYGGPGCK"
     exon            8685. .8915
                     /gene="F24G24.20"
                     /number=1
     intron          8916. .9401
                     /gene="F24G24.20"
                     /number=1
     exon            9402. .9554
                     /gene="F24G24.20"
                     /number=2
     intron          9555. .10284
                     /gene="F24G24.20"
                     /number=2
     exon            10285. .10641
                     /gene="F24G24.20"
                     /number=3
     gene            11823. .12747
                     /gene="F24G24.30"
     CDS             join(11823. .11915,12019. .12747)
                     /gene="F24G24.30"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB39776.1"
                     /db_xref="GI:4538952"
                     /translation="MTSILSAVSKLRGCVNQIENKNPSGASEEDILNQAKMLLTQYEK
                     YKRGFKFDHVWPILKGIEKFANDNMKTPPAFQGEGRDVTSSSSFSINTESSPSPGMNS
                     TDLNMDSEDANFSLSSRPKGLKKAKRKQQSEEQFKQLLEQNDKLIKAITKGTSLNEI
                     QRQKIEVARMKEENKILFADLNSISDPSSRAYVENERKRILEKRAQTNQHEEDGESQ
                     YHGSQYRASHYQESLFHGKQVQGEPDQGEDKRSPNNQEDFTQYYNYLSGTGNNFP"
     exon            11823. .11915
                     /gene="F24G24.30"
                     /number=1
     intron          11916. .12018
                     /gene="F24G24.30"
                     /number=1
     exon            12019. .12747
                     /gene="F24G24.30"
                     /number=2
     gene            14529. .15088
                     /gene="F24G24.40"
     gene            complement(join(14529. .14816,14888. .15088))
                     /gene="F24G24.40"
     CDS             complement(join(14529. .14816,14888. .15088))
                     /gene="F24G24.40"
                     /note="strong similarity to zinc-finger protein R2931,
                     Oryza sativa, PIR3:JE0116"
                     /codon_start=1
                     /product="zinc-finger-like protein"
                     /protein_id="CAB39777.1"
```

```
                /db_xref="GI:4538953"
                /translation="MKIQCEVCEKAEAFVLCCSDEAVLCKPCDIKVHEANKLFQPHHR
                VALQKDAASATTASGAPLCDICQERKGYFFCLEDRAMLCNDCDEAIHTCNSHQRFLLS
                GVQVSDQSLTENSECSTSFSSETYQIQSKVSLNSQYSSEETEAGNSGEIVHKNPSVIL
                SP"
     on         complement(14529. .14816)
                /gene="F24G24.40"
                /number=1
     tron       complement(14817. .14887)
                /gene="F24G24.40"
                /number=1
     on         complement(14888. .15088)
                /gene="F24G24.40"
                /number=2
     ne         16099. .16686
                /gene="F24G24.50"
     S          16099. .16686
                /gene="F24G24.50"
                /note="contains EST gb:AA597970"
                /codon_start=1
                /product="heat shock protein 22.0"
                /protein_id="CAB39778.1"
                /db_xref="GI:4538954"
                /translation="MMKHLLSIFFIGALLLGNIKTSEGSLSSALETTPGSLLSDLWLD
                RFPDPFKILERIPLGLERDTSVALSPARVDWKETAEGHEIMLDIPGLKKDEVKIEVEE
                NGVLRVSGERKREEEKKGDQWHRVERSYGKFWRQFKLPDNVDMESVKAKLENGVLTIN
                LTKLSPEKVKGPRVVNIAAEEDQTAKISSSESKEL"
     on         16099. .16686
                /gene="F24G24.50"
                /number=1
     ne         17048. .18332
                /gene="F24G24.60"
     ne         complement(join(17048. .17182,17267. .17393,17462. .17596,
                17671. .18081,18156. .18332))
                /gene="F24G24.60"
     S          complement(join(17048. .17182,17267. .17393,17462. .17596,
                17671. .18081,18156. .18332))
                /gene="F24G24.60"
                /EC_number="2.7.1.4"
                /note="strong similarity to fructokinase, Lycopersicon
                esculentum, gb:U62329
                Contains pfkB family of carbohydrate kinases signatures,
                Pfkb_Kinases_1 [GGAPANVACAITKLGGKSAFIGKFG], Pfkb_Kinases_2
                [DTTGAGDSFVGAFL]"
                /codon_start=1
                /product="fructokinase-like protein"
                /protein_id="CAB39779.1"
                /db_xref="GI:4538955"
                /translation="MANTPLIVSFGEMLIDFVPDTSGVSLAESTGFLKAPGGAPANVA
                CAITKLGGKSAFIGKFGDDEFGHMLVNILKKNGVNSEGVCFDTNARTALAFVTLKKDG

  Que  Match            14.4%;  Score 113.8;  DB 8;  Length 93856;
  Bes  Local Similarity  70.4%;  Pred. No. 3e-19;
  Mat  es  171;  Conservative    0;  Mismatches   62;  Indels   10;  Gaps    1;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           |||||||||||| ||||    ||  || |  ||  |||  | ||||||
Db    1  98 ATCATCAACAAACACAAAATCGAAACCAAATCACTTAAACAAGAAAGAA----------A 19049

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
             |||| | ||||| ||| |||||||| | ||||  || || |  ||||||||||||||
Db    1  48 CAAAGAGAAATGAGCTCTGCAAGCAAAACGTGGATGGTTGCTGCAAGCATCGGAGCCGTT 18989

Qy      21 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           |||||||| ||||||||||||||   | ||||| |||||||| |  |||| || |  |||
Db    1  88 GAGGCATTGAAAGACCAACTAGGCGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAAT 18929

Qy      81 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           ||  |||| || |||||| | ||||| || || ||||  || | | |||||| || ||
Db    1 -28 CAGTATCTACGCAACAACTTAAGATCCGTGTCGCAAGCTAAGAAGCTCTCTTCCTCATCA 18869

Qy     241 GTC 243
           |||
```

```
Db  18868 ATC 18866


RESULT 12
ATCHRIV29/c
LOCUS       ATCHRIV29             199861 bp    DNA     linear   PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 29.
ACCESSION   AL161517
VERSION     AL161517.2  GI:7267723
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 155039)
  AUTHORS   Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 147497 to 199861)
  AUTHORS   Wedler,H., Wedler,E., Wambutt,R., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 199861)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV28 at the 5' end and an
            overlap with ATCHRIV30 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .199861
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     gene            7035. .7307
                     /gene="AT4g10270"
     CDS             7035. .7307
                     /gene="AT4g10270"
                     /note="strong similarity to wound-induced protein,
                     Lycopersicon esculentum, PIR2:S19773
                     contains EST gb:AI995575.1"
                     /codon_start=1
                     /product="probable wound-induced protein"
                     /protein_id="CAB78150.1"
                     /db_xref="GI:7267724"
                     /translation="MSSTSKAWTVAVSIGAVEALKDQLGLCRWNYILRSVNQHLRNNV
                     RSVSQGKRFSSSSVSAAVTSSGESEKAKKAEESLRTVMYLSCWGPN"
     exon            7035. .7307
                     /gene="AT4g10270"
                     /number=1
     gene            7450. .8226
                     /gene="AT4g10280"
     gene            complement(join(7450. .7666,8021. .8226))
                     /gene="AT4g10280"
     CDS             complement(join(7450. .7666,8021. .8226))
                     /gene="AT4g10280"
                     /note="similarity to predicted protein, Arabidopsis
                     thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB78151.1"
                     /db_xref="GI:7267725"
                     /translation="MGNHFLTLSLLLVTVCVCVSFITTKLNPKEAIVSVSSDSEIPTE
                     IHGVKILRQASDTKLAQLGVASWPKWEGAPSKFPWEFKKTETIYFMEGKVKVNVDGYD
```

```
                EEEETPEIGKGDVVVPPKDMKVVWEITEAVKKQYSIEF"
exon            complement(7450. .7666)
                /gene="AT4g10280"
                /number=1
intron          complement(7667. .8020)
                /gene="AT4g10280"
                /number=1
exon            complement(8021. .8226)
                /gene="AT4g10280"
                /number=2
tRNA            10468. .10619
                /product="tRNA-Ser"
                /note="codon recognized: GGA"
gene            12188. .12418
                /gene="AT4g10290"
gene            complement(12188. .12418)
                /gene="AT4g10290"
CDS             complement(12188. .12418)
                /gene="AT4g10290"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAB78152.1"
                /db_xref="GI:7267726"
                /translation="MPRKFPWKFKKTETMYFVEGKLKVKVEDHHKEGEALEFVAGDLV
                VFPQDMNVFVDVIEDVKKRYYRESEIEESELP"
exon            complement(12188. .12418)
                /gene="AT4g10290"
                /number=1
gene            16794. .18176
                /gene="AT4g10300"
CDS             join(16794. .16993,17972. .18176)
                /gene="AT4g10300"
                /note="similarity to predicted protein, Arabidopsis
                thaliana
                contains EST gb:AI823169.1"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAB78153.1"
                /db_xref="GI:7267727"
                /translation="MGCIGVVNVTSINPYLTRRSNKPYNSRRPSSMAAAIRAESTEKL
                GITIEKNPPESKLTQLGVRSWPKWGCPPSKFPWTYSAKETCYLLQGKVKVYPNGSDEG
                VEIEAGDFVVFPKGMSCTWDVSVAVDKHYQFE"
exon            16794. .16993
                /gene="AT4g10300"
                /number=1
intron          16994. .17971
                /gene="AT4g10300"
                /number=1
exon            17972. .18176
                /gene="AT4g10300"
                /number=2
gene            24238. .27814
                /gene="AT4g10310"
CDS             join(24238. .25310,26586. .26819,27652. .27814)
                /gene="AT4g10310"
                /note="similarity to potassium transport protein TRK1,
                Saccharomyces cerevisiae, PIR2:JU0466"
                /codon_start=1
                /product="potassium uptake transporter-like protein"
                /protein_id="CAB78154.1"
                /db_xref="GI:7267728"
                /translation="MDRVVAKIAKIRSQLTKLRSLFFLYFIYFLFFSFLGFLALKITK
                PRTTSRPHDFDLFFTSVSAITVSSMSTVDMEVFSNTQLIFLTILMFLGGEIFTSFLNL
                YVSYFTKFVFPHNKIRHILGSYNSDSSIEDRCDVETVTDYREGLIKIDERASKCLYSV
                VLSYHLVTNLVGSVLLLVYVNFVKTARDVLSSKEISPLTFSVFTTVSTFANCGFVPTN
                ENMIIFRKNSGLIWLLIPQVLMGNTLFPCFLVLLIWGLYKITKRDEYGYILKNHNKMG
                YSHLLSVRLCVLLGVTVLGFLIIQLLFFCAFEWTSESLEGMSSYEKLVGSLFQVVNSR
                HTGETIVDLSTLSPAILVLFILMMYLPPYTLFMPLTEQKTIEKEGGDDDSENGKKVKK
                SGLIVSQLSFLTICIFLISITERQNLQRDPINFNVLNITLEVIRIIINKISFVMAVHME
                TLVSLPGTAVNGVWTSAMVAAKTRVMGLQDDGVQWENSY"
exon            24238. .25310
                /gene="AT4g10310"
                /number=1
intron          25311. .26585
                /gene="AT4g10310"
                /number=1
exon            26586. .26819
                /gene="AT4g10310"
                /number=2
intron          26820. .27651
                /gene="AT4g10310"
                /number=2
exon            27652. .27814
                /gene="AT4g10310"
                /number=3
gene            29756. .36739
                /gene="AT4g10320"
gene            complement(join(29756. .30244,30328. .30393,30579. .30701,
                30870. .30937,31013. .31076,31128. .31311,31418. .31506,
                31653. .31716,31808. .31911,32003. .32140,32227. .32352,
                32462. .32671,32770. .32853,32977. .33136,33651. .33784,
                33863. .33952,34221. .34379,34479. .34557,34914. .35011,
                35156. .35269,35441. .35962,36051. .36116,36206. .36739))
                /gene="AT4g10320"
CDS             complement(join(29756. .30244,30328. .30393,30579. .30701,
                30870. .30937,31013. .31076,31128. .31311,31418. .31506,
                31653. .31716,31808. .31911,32003. .32140,32227. .32352,
                32462. .32671,32770. .32853,32977. .33136,33651. .33784,
                33863. .33952,34221. .34379,34479. .34557,34914. .35011,
                35156. .35269,35441. .35962,36051. .36116,36206. .36739))
                /gene="AT4g10320"
                /note="strong similarity to isoleucine--tRNA ligase, Homo
                sapiens, PIR2:I59314
                Contains Aminoacyl-transfer RNA synthetases class-I
                signature AA49-60;Prokaryotic membrane lipoprotein lipid
                attachment site AA789-798;Prokaryotic membrane lipoprotein
                lipid attachment site AA1215-1225"
                /codon_start=1
                /product="isoleucine-tRNA ligase-like protein"
                /protein_id="CAB78155.1"
                /db_xref="GI:7267729"
                /translation="MEEVKEGKEFSFPRQEEDVLSFWTEIDAFKTQLKRTENLPEYIF
                YDGPPFATGLPHYGHILAGTIKDIVTRYQTMTGHHVTRRFGWDCHGLPVENEIDRKLN
                IKRRDEVIKMGIDKYNEECRSIVTRYVAEWEKVITRCGRWIDFKNDYKTMDLPFMESV
                MWVFSQLWEKNLVYRGFKVMPYSTGCKTPLSNFEAGQNYKEVPDPEIMVTFPVIGDQD
                NAAFVAWTTTPWTLPSNLALCVNAKFVYVKVRNKNTGKVYIVAESRLSALPTDKPKAK
                LSNGPAGDTKKANPKAKGAKPESAADSYEVLEKFNGASLVGKKYEPLFDYFSDFSSEA
                FRVVADDYVTDDSGTGIVHCAPAFGEDDYRVCLLNKIIKKGENLVVAVDDDGLFTERI
                THFSGRYVKDADKDIIEAVKAKGRLVKTGSFTHSYPFCWRSDTPLIYRAVPSWFVRVE
                LKEKLLKSNEQTEWVPGYVKDKRFHNWLENARDWAISRSRFWGTPLPIWISDDGEEV
                VTMDSVEKLEKLSGVKVFDLHRHHIDHITIPSSRGDEFGVLRRVEDVFDCWFESGSMP
                YAYIHYPFENKELFEKNFPADFVAEGLDQTRGWFYTLMVLSTALFEKPAFKNLICNGL
                VLAEDGKKMAKKLRNYPPDLEVIDEYGADAVRLYLINSPVVRAEPLRFKKEGVLGVVK
                DVFLPHYNAYRFLVQNAKLETEGGVPFVPTDLATIQSANILDQWIHSATCSLVRFV
                EEMDAYRLYIVVPRLLKFLDNLTNIYVRFNRKRLKGRTGEDDCHTALSTLFNVLLTSI
                KVMAPFTPFFTETLYQNLRKACKGSEESVHYCSIPPREGMEGERIELSVTRMMKIIDI
                ARNIRERNKLPLKTPLKEITVVHPDADFLNDITGVLREYVLEELNVRSLVPCNDTLS
                ASLKAEPDFSVLGKRLGKSMGLVAKEVKEMSQKDILAFEEAGEVTIANHLLKETDIFV
                SHAMMFSCEIALRIVRVFKRPDDLKENEIDSAGDGDVLVILDLRADDSLVEAGFARE
                VIRIQKLRKKSGLEPTDFVEVYFQSLDEDESVSKQVLVSQEQNIKDSIGSTLLLSSLN
                PSHAVIIADETFTPKETSDESVKKVPKLSYKISLARPALKFNEEAVLALYSGESPSHH
                NRFSLSILSRFLFNVYIYGSRRCEICNRASNILVIKGSFEFEVRVPSRRWEGSLVKC
```

```
Query Match          14.4%;  Score 113.8;  DB 8;  Length 199861;
Best Local Similarity  70.4%;  Pred. No. 3.3e-19;
Matches  171;  Conservative    0;  Mismatches   62;  Indels   10;  Gaps    1;

Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           ||||||||||||| ||||     ||  || |  ||  |||  | ||||||
Db    5766 ATCATCAACAAACACAAAATCGAAACCAAATCACTTAAACAAGAAAGAA----------A 5717

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
           |||| | ||.|| ||| |||||||| | ||||  || || |  ||||||||||||||
Db    5716 CAAAGAGAAATGAGCTCTGCAAGCAAAACGTGGATGGTTGCTGCAAGCATCGGAGCCGTT 5657

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
```

```
            |||||||| |||||||||||||| | ||||| ||||||||| | |||| || | |||
Db    5656 GAGGCATTGAAAGACCAACTAGGCGTGTGTCGTTGGAACTACGTGATCCGATCTGCGAAT 5597

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTCT 240
           || |||| || ||||||| | ||||| || || ||||  || | | ||||||| || ||
Db    5596 CAGTATCTACGCAACAACTTAAGATCCGTGTCGCAAGCTAAGAAGCTCTCTTCCTCATCA 5537

Qy     241 GTC 243
           ||
Db    5536 ATC 5534
```

RESULT 13
AC130811
LOCUS AC130811 141382 bp DNA linear HTG 23-AUG-2002
DEFINITION Medicago truncatula clone mth2-26b4, WORKING DRAFT SEQUENCE, 9 unordered pieces.
ACCESSION AC130811
VERSION AC130811.4 GI:22450619
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 141382)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-26b4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 141382)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE 3 (bases 1 to 141382)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT On Aug 23, 2002 this sequence version replaced gi:22297364.
---------- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
---
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2895: contig of 2895 bp in length
* 2896 2995: gap of unknown length
* 2996 7176: contig of 4181 bp in length
* 7177 7276: gap of unknown length
* 7277 13609: contig of 6333 bp in length
* 13610 13709: gap of unknown length
* 13710 19986: contig of 6277 bp in length
* 19987 20086: gap of unknown length
* 20087 36269: contig of 16183 bp in length
* 36270 36369: gap of unknown length
* 36370 60069: contig of 23700 bp in length
* 60070 60169: gap of unknown length
* 60170 85493: contig of 25324 bp in length
* 85494 85593: gap of unknown length
* 85594 112631: contig of 27038 bp in length
* 112632 112731: gap of unknown length
* 112732 141382: contig of 28651 bp in length.
FEATURES Location/Qualifiers
source 1. .141382
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="mth2-26b4"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 47478 a 23124 c 22805 g 47171 t 804 others
ORIGIN

Query Match 11.4%; Score 90.2; DB 2; Length 141382;
Best Local Similarity 61.1%; Pred. No. 6.4e-13;
Matches 146; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```
Qy       1 ATCATCAACAAAAACAATTCTCAATACACAAAACACAAAACACAAAGAAGTTTAATTCTC 60
           |||||||   |  | ||       |    ||  | |  |  ||  ||        | ||
Db   49239 ATCATCACACATCAAAACAAAACAACTTCATCATAATATTCAACAAAGGTACAAGTTACA 49298

Qy      61 TGAAGAAAGATGAGTTCTACAAGCAAAGCATGGACAGTGGCAGTGAGCATCGGAGCCGTA 120
             ||  || |||||| |  |||  ||||| |||  ||| |||   ||||| ||||| ||
Db   49299 ACAAACAACATGAGTGCAGCAACAAAAGCTTGGGTAGTAGCATCAAGCATAGGAGCAGTG 49358

Qy     121 GAGGCATTAAAAGACCAACTAGGTCTTTGTCGGTGGAACTACATACTCCGGTCGGTTAAT 180
           || |||||||||||||||| ||| | ||  ||||||||||      |  | ||  || ||
Db   49359 GAAGCATTAAAAGACCAACTTGGTGTATGCAGGTGGAACTATGCTTTTAGATCACTTCAT 49418

Qy     181 CAACATCTCCGGAACAACGTTAGATCTGTTTCTCAAGGGAAAAGGTTCTCTTCGTCTTC 239
           ||||||    ||||||  |||||||   |||||||  || |  || || ||  ||||
Db   49419 CAACATGCTAAGAACAATATTAGATCATACTCTCAAGCCAAGAAATTGTCCTCTGCTTC 49477
```

RESULT 14
AF153284
LOCUS AF153284 847 bp mRNA linear PLN 02-JUN-1999
DEFINITION Arabidopsis thaliana putative progesterone-binding protein homolog (Atmp1) mRNA, complete cds.
ACCESSION AF153284
VERSION AF153284.1 GI:4960155
KEYWORDS .
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 847)
AUTHORS Choi,J.H., Choi,H. and Gray,P.
TITLE Plant homologues of mammalian putative progesterone-binding membrane proteins
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 847)
AUTHORS Choi,J.H., Choi,H. and Gray,P.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-1999) School of Biology, Georgia Institute of Technology, MC 0230, Atlanta, GA 30332-0230, USA
FEATURES Location/Qualifiers
source 1. .847
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="dbEST:N37902"
/db_xref="taxon:3702"
/chromosome="5"
gene 1. .847
/gene="Atmp1"
CDS 62. .724
/gene="Atmp1"
/codon_start=1
/product="putative progesterone-binding protein homolog"
/protein_id="AAD34616.1"
/db_xref="GI:4960156"
/translation="MALELWQTLKEAIHAYTGLSPVVFFTALALAFAIYQVISGWFAS
PFDDVNRHQRARSLAQEEEPPIPQPVQVGEITEEELKQYDGSDPQKPLLMAIKHQIYD
```
```

VTQSRMFYGPGGPYALFAGKDASRALAKMSFEEKDLTWDISGLGPFELDALQDWEYKF
MSKYAKVGTVKVAGSEPETASVSEPTENVEQDAHVTTTPEKTVVDKSDDAPAETVLKK
EE"
BASE COUNT 237 a 185 c 192 g 233 t
ORIGIN

Query Match 11.4%; Score 89.6; DB 8; Length 847;
Best Local Similarity 56.6%; Pred. No. 4.9e-13;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
Db    281 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAGCCCCTTCTT 340

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
Db    341 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTACGGACCAGGA 400

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
Db    401 GGGCCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAGATGTCATTT 460

Qy    517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTT 576
Db    461 GAGGAGAAAGACTTGACTTGGGATATCTCTGGTCTTGGTCCCTTTGAGCTAGATGCTCTT 520

Qy    577 AATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT 626
Db    521 CAAGATTGGGAGTACAAGTTCATGAGCAAGTATGCTAAGGTTGGTACTGT 570
```

RESULT 15
AY086811
LOCUS AY086811 889 bp mRNA linear PLN 25-JUN-2002
DEFINITION Arabidopsis thaliana clone 28003 mRNA, complete sequence.
ACCESSION AY086811
VERSION AY086811.1 GI:21405521
KEYWORDS FLI_CDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 889)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. (2002) In press
REFERENCE 2 (bases 1 to 889)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-Length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 889)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or LAer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0.

conset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES Location/Qualifiers
source 1. .889
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="28003"
CDS 45. .707
/codon_start=1
/product="progesterone-binding protein-like"
/protein_id="AAM63860.1"
/db_xref="GI:21555444"
/translation="MALELWQTLKEAIHAYTGLSPVVFFTALALAFAIYQVISGWPAS
PFDDVNRHQRARSLAQEEEPPIPQPVQVGEITEEELKQYDGSDPQKPLLMAIKHQIYD
VTQSRMFYGPGGPYALFAGKDASRALAKMSFEEKDLTWDISGLGPFELDALQDWEYKP
MSKYAKVGTVKVAGSEPETASVSEPTENVEQDAHVTTTPEKTVVDKSDDAPAETVLKK
EE"
BASE COUNT 249 a 188 c 202 g 250 t
ORIGIN

Query Match 11.4%; Score 89.6; DB 8; Length 889;
Best Local Similarity 56.6%; Pred. No. 4.9e-13;
Matches 164; Conservative 1; Mismatches 125; Indels 0; Gaps 0;

```
Qy    337 GAATTCACCGCAGAGCAGCTAAGCCAATACAACGGCACCGACGAATCAAAGCCGATCTAC 396
Db    264 GAGATCACGGAGGAGGAGCTTAAACAGTACGATGGCTCTGATCCTCAAAAGCCCCTTCTT 323

Qy    397 GTCGCAATCAAAGGCCGTGTGTTCGAYGTCACCACCGGAAAATCCTTCTACGGCTCCGGA 456
Db    324 ATGGCTATCAAACATCAGATCTATGATGTTACACAAAGCAGGATGTTCTACGGACCAGGA 383

Qy    457 GGCGATTACTCGATGTTCGCCGGAAAAGACGCGAGCAGAGCTTTGGGTAAGATGAGTAAG 516
Db    384 GGACCATATGCTTTGTTTGCAGGGAAAGACGCTAGCCGAGCTCTTGCAAAGATGTCATTT 443

Qy    517 AACGAAGAAGATGTGTCTCCTTCTCTTGAAGGTCTCACTGAGAAAGAGATCAATACTCTT 576
Db    444 GAGGAGAAAGACTTGACTTGGGATATCTCTGGTCTTGGTCCCTTTGAGCTAGATGCTCTT 503

Qy    577 AATGATTGGGAGACCAAATTTGAAGCTAAGTATCCTGTCGTTGGCCGTGT 626
Db    504 CAAGATTGGGAGTACAAGTTCATGAGCAAGTATGCTAAGGTTGGTACTGT 553
```

Search completed: January 8, 2003, 16:28:28
Job time : 3899 secs

Nucleic Acids Research

OXFORD Journals online

 4139-4146

Abstract of this Article
Reprint (PDF) Version of this Article
Similar articles found in:
Nucl. Acids. Res. Online
PubMed
PubMed Citation
Search Medline for articles by:
Uddin, A. H. || Krull, U. J.
Alert me when:
new articles cite this article
Download to Citation Manager

A fiber optic biosensor for fluorimetric detection of triple-helical DNA

André H. Uddin, Paul A. E. Piunno[1], Robert H. E. Hudson[1,+], Masad J. Damha* and Ulrich J. Krull[1,*]

Department of Chemistry, Otto Maas Chemistry Building, McGill University, Montreal, Quebec H3A 2K6, Canada and [1]Chemical Sensors Group, Department of Chemistry, Erindale College, University of Toronto, Mississauga, Ontario L5L 1C6, Canada

Received May 13, 1997; Revised and Accepted August 22, 1997

ABSTRACT

A fiber optic biosensor was used for the fluorimetric detection of T/AT triple-helical DNA formation. The surfaces of two sets of fused silica optical fibers were functionalized with hexaethylene oxide linkers from which decaadenylic acid oligonucleotides were grown in the 3' to 5' and 5' to 3' direction, respectively, using a DNA synthesizer. Fluorescence studies of hybridization showed unequivocal hybridization between oligomers immobilized on the fibers and complementary oligonucleotides from the solution phase, as detected by fluorescence from intercalated ethidium bromide. The complementary oligonucleotide, dT_{10}, which was expected to Watson-Crick hybridize upon cooling the system below the duplex melting temperature (*T*m), provided a fluorescence intensity with a negative temperature coefficient. Upon further cooling, to the point where the pyrimidine motif T*AT triple-helix formation occurred, a fluorescence intensity change with a positive temperature coefficient was observed. The reverse-Hoogsteen T·AT triplex, which is known to form with branched nucleic acids, provided a corresponding decrease in fluorescence intensity with decreasing temperature. Full analytical signal evolution was attainable in minutes.

INTRODUCTION

With recent advances in nanotechnology (1), there is an increased demand to investigate biomolecular structure and behavior (2). One particular area of interest stems from the progress in the synthesis of novel nucleic acid macromolecules. Dendrimers (3,4), circular (5) and cage oligonucleotides (6) have been synthesized and these novel compounds are finding applications in biotechnology (7,8).

Furthermore, there is much interest in the development of devices for rapid diagnostic assays to detect microorganisms, viruses and genetic mutations based on hybridization with immobilized nucleic acid probes. Approaches involving electrochemical (9), acoustic wave or piezoelectric (10), plasmon resonance (11,12), colorimetric sensing of non-particle aggregates (13) and fluorescence based optical fiber sensing techniques have been proposed (14-16). In these examples, identification of the analyte is based on the occurrence of Watson-Crick hybridization events, with the formation of three-stranded structures, or triplexes, being largely ignored.

Triple-helical oligonucleotides have potential use as sequence specific artificial nucleases (17), modulators of DNA-binding proteins/gene expression (18,19; for a recent review see ref. 20), materials for genomic mapping (21), and sensitive screening reagents to detect mutations within duplex DNA (22). Formation of three-stranded helices by nucleic acids is a well-known phenomenon which involves a third strand interacting with a purine rich strand in the underlying Watson-Crick DNA duplex (23,24). Two distinct classes of DNA triple-helices have been characterized which differ in the composition and orientation of the third strand relative to the Hoogsteen binding partner (25-32). Nucleic acid binding ligands can be used to identify DNA structures and morphology. For example, ethidium bromide binds to both duplexes and triplexes by intercalation (33), but there is a marked difference in the binding efficiency and fluorescence quantum efficiency between both types of complexes (34-36).

We have focused on the use of a nucleic acid binding ligand (e.g., ethidium bromide) and fluorescence transduction strategy to investigate oligonucleotide hybridization on fused silica optical fiber surfaces. Previously, we reported detection of hybridization events between fibers derivatized with single-stranded deoxyribonucleic acid and complementary DNA and RNA from solution (14). Herein, we report the use of optical biosensor technology for rapid detection of T/AT triplex formation in both parallel and antiparallel configurations. This rapid and efficient triple-helical assay may be extended to include diagnostic assays for sequence-specific duplex recognition, monitoring *in vivo* concentration of gene therapy pharmaceuticals, and for studying properties of synthetic oligonucleotides.

MATERIALS AND METHODS

Chemicals

Reagent grade solvents were purchased (BDH, Toronto, ON) and further purified or dried by standard laboratory practices. DNA synthesis reagents and decadeoxyadenylate (dA_{10}) were purchased from Dalton Chemical Laboratories Inc. (Toronto, ON) and were used as received or were prepared as below. Anhydrous acetonitrile (Dalton) was predried by distillation from P_2O_5 and redistilled from calcium hydride under dry argon. Tetrahydrofuran (BDH) was predried over CaH_2, filtered and distilled immediately prior to use from sodium metal (Aldrich)/benzophenone (Aldrich). Ethidium bromide (3,8-diamino- 5-ethyl-6-phenylphenanthridinium bromide; Aldrich) was used as received. Water was double-distilled in glass, treated with diethyl pyrocarbonate (Aldrich) and autoclaved. Molecular biology grade polyacrylamide gel electrophoresis reagents and apparatus were obtained through Bio-Rad (Hercules, CA). Silica gel (Toronto Research Chemicals, Toronto, ON) had a particle size of 30-70 microns. Pre-cut fused silica optical fiber pieces with a length of 48 mm and a core diameter of 400 μm having both termini polished to within a 0.3 μm tolerance were obtained from 3M Specialty Optical Fiber (North York, Ontario, Canada) in addition to lengths of fiber having the same core material and diameter with a TECS 48 low refractive index outer cladding (0.48 numerical aperture).

Derivatization of optical fibers

Synthesis of DMT-HEG (dimethoxytrityled hexaethylene glycol). A solution of dimethoxytrityl chloride (7.1 g, 21 mmol) in dry pyridine (10 ml) was added dropwise to a stirred solution of hexaethylene glycol (HEG, 5.6 ml, 21 mmol in 5 ml pyridine) under an argon atmosphere. Stirring was continued overnight after which the reaction mixture was combined with dichloromethane (50 ml). The mixture was then shaken with 5% aqueous bicarbonate (2 × 90 ml) and then with water (2 × 90 ml) to remove unreacted HEG, pyridine and salts. The organic layer was dried under reduced pressure to yield the crude product. The product was purified by silica gel column chromatography using an eluent of 1:1 dichloromethane/diethyl ether containing 0.1% triethylamine (2.9g, 24% yield). The identity of the product was confirmed by proton NMR spectroscopy. R_f (silica gel thin-layer chromatography): 0.10 in CH_2Cl_2/ether (1:1). ^1H NMR (200 MHz, $CDCl_3$) [delta]: 7.48 (t, 1H, J = 1.8 Hz), 7.46-7.42 (m, 2H), 7.27 (d, 1H, J = 2.6), 7.3 (d, 1H, J = 3.3 Hz) , 7.1 (m, 8H), 3.79 (s, 6H), 3.64 (s, 24H).*Surface preparation of optical fibers.* The coating material was mechanically stripped from the pre-cut optical fiber pieces and the cladding dissolved by treatment with acetone. The surface of the fibers were then cleaned via treatment with 25% ammonia/30% hydrogen peroxide/water (1:1:5, v/v/v) for 5 min at 80°C followed by rinsing with 30% hydrogen peroxide. The fibers were then treated with a solution of conc. HCl/30% hydrogen peroxide/water (1:1:5, v/v/v) for 5 min at 80°C, followed by rinsing with methanol, dichloromethane and diethyl ether.*Functionalization of optical fibers with 3-glycidopropyltrimethoxysilane (GOPS).* Following a modification of the method reported by Maskos and Southern (37), optical fibers and silica gel were activated by placement into a solution of xylene/GOPS/diisopropylethylamine (100:30:1 v/v/v). The reaction was permitted to proceed with gentle agitation for 24 h under nitrogen at 80°C. The fibers and silica gel were rinsed with methanol, dichloromethane and diethyl ether. *Linkage of DMT-HEG to GOPS functionalized optical fibers.* The fibers and silica gel were then functionalized with monotritylated hexaethylene glycol (DMT-HEG) (250 mg, 0.46 mmol) in 30 ml of xylene containing a catalytic amount of sodium hydride with gentle agitation at 40°C. Silica gel samples (~10 mg) were taken from the reaction mixture daily to determine the loading of DMT-HEG, and this was presumed to indicate loading on the activated fibers. The silica gel samples were immediately washed with 10 ml portions of dichloromethane until the wash solution showed no absorption at 504 nm upon treatment with trichloroacetic acid. The GOPS-HEG-DMT functionalized silica gel samples were then dried under reduced pressure and treated with 5 ml of 5% trichloroacetic acid in dichloroethane in order to liberate the dimethoxytrityl moieties from the hexaethylene glycol chains. The absorbance (504 nm) of the resulting solution was then determined to quantitatively measure the loading of immobilized DMT-HEG. This analysis indicated that the reaction had gone to completion after 7 days. After this time, the fibers were removed from the reaction mixture, washed with dichloromethane and dried by storage *in vacuo* and over P_2O_5 overnight.

The secondary hydroxyl groups produced after reaction of the HEG linker with the epoxide moieties and all other silanols were capped via treatment with trimethylsilyl chloride in pyridine (1:10 v/v) under argon at room temperature for 16 h followed by treatment with acetic anhydride/*N*-methylimidazole/collidine in THF to prevent unwanted oligonucleotide growth at these sites (38). The fibers were then washed sequentially with pyridine, dichloromethane, methanol and diethyl ether and stored *in vacuo* and over P_2O_5. The amount (or 'loading') of DMT-HEG spacers on the surface of a fused silica fiber was ~1 nmol/fiber (48 mm in length).

Synthesis of oligonucleotides on optical fibers

Approximately 10 functionalized DMT-HEG-GOPS fibers (48 mm in length) were placed in a standard 10 μmol scale Applied Biosystems synthesis column and capped with acetic anhydride prior to DNA synthesis using the ABI supplied cycle. Detritylation was performed with 3% trichloroacetic acid in

dichloroethane. Activation of phosphoramidites was achieved with 0.5 M tetrazole in acetonitrile. Reagents for capping were as follows: Cap A, phenoxyacetyl anhydride Cap A reagent from Millipore (Mississauga, ON); and Cap B, 16% *N*-methylimidazole in THF (w/v). Iodine, 0.1 M, in THF/pyridine/water (25:20:2, v/v/v) was used for oxidations. Phenoxyacetyl protected dG, dC, dA phosphoramidite monomers were obtained from Millipore. N^6-phenoxyacetyl-3'-*O*-DMT-2'-deoxyadenosine-5'-*O*-[([beta]-cyanoethyl)*N*,*N*-diisopropyl]-phosphorami was prepared via standard protocols (39). The oligomers were deprotected with conc. NH_4OH solution for 2 h at room temperature. Following deprotection, the ammonia solution was collected, the column was washed with autoclaved water and the eluent was also kept. Quantitation of the eluents at 260 nm indicated that ~20% of the oligomers remained bound to the fiber surface.

Synthesis of branched oligonucleotides

The 'V' branched sequence **1** (Fig. 3) was synthesized on an Applied Biosystems 381A instrument using a 1 μmol scale synthesis cycle and [beta]-cyanoethylphosphoramidite chemistry (3,40). Purification, desalting, and analysis of the branched oligonucleotide **1** was accomplished by our detailed protocols (3,41). Typical isolated yields of this branched oligomer were 15-25% (~0.4-1.5 mg), as determined by absorption at 260 nm.

UV thermal denaturation and renaturation studies

Absorbance versus temperature profiles of the nucleic acid complexes (10 mM Tris, 50 mM $MgCl_2$, pH 7.3, 2.5×10^{-8} M ethidium bromide) were measured at 260 nm using a Varian Cary I UV-VIS spectrophotometer. Thermal denaturation profiles (i.e., melting curves) and thermal renaturation profiles (i.e., cooling curves) of each system of oligonucleotides were acquired at two temperature ramp rates, 0.5°C/min and 0.06°C/min. For each system of oligonucleotides, the denaturation and renaturation profiles provided identical results for the melting temperature (T_m) and showed no dependence on the temperature ramp rate used. Normalized plots were constructed according to the method of Kibler-Herzog *et al.* (42). All complexes showed sharp melting transitions. The values of T_m were determined from the first derivative of each thermal curve with an error in precision not greater than ±0.5°C based on variance in repeated experiments.

Figure 1. Schematic diagram of the apparatus used for fluorescence investigations of nucleic acid hybridization on the fiber optic sensor.

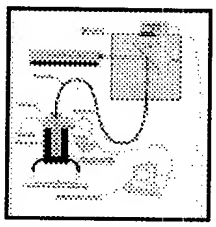


Figure 2. Derivatization of fused silica optical fibers.

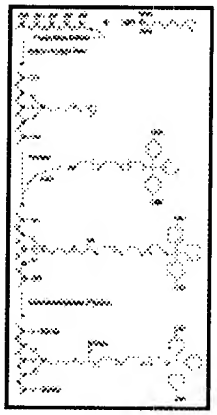
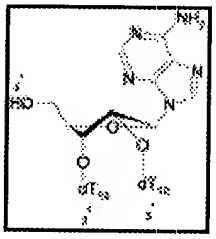


Figure 3. The chemical structure of compound **1**, a branched oligonucleotide with identical chains linked to the 2'- and 3'-positions of a ribose branch-point nucleoside, i.e., rA[(2'-5'-dT_{10})/(3'-5'-dT_{10})]. Ad, adenosine; Th, thymine. Two molecules of dT_{10} hybridize with dA_{10} to give the more common parallel (T*AT, Hoogsteen) triplex, whereas **1** forms a triplex in an antiparallel (T·AT, reversed-Hoogsteen) binding motif.



Instrument setup and fluorescent measurements

The instrument used for fluorescence intensity measurements was based on a fluorescence microscope as was previously described by Krull and co-workers (43) and shown in Figure 1. Radiation from an Ar^+ laser operated at 488 nm was reflected by the dichroic mirror (495 nm cut-off) in the fluorescence microscope to a Zeiss 16* immersion lens with a numerical aperture of 0.5 (Empix Imaging, Mississauga, ON, Canada). The laser radiation exciting the immersion lens was coupled into a delivery fiber of similar numerical aperture (0.48) aligned beneath the objective. The light was totally internally reflected along the length of the delivery fiber to a sensing fiber functionalized with immobilized oligonucleotide. Coupling of the radiation between fibers was achieved by abutting the distal terminus of the delivery fiber to the proximal terminus of the sensing fiber. A loss in optical transmission of no more than 2% was observed for the coupled system. The termini of the teflon fiber coupler were designed as compression-fit ends which provided a solution-tight seal that prevented contaminants from diffusing into the fiber coupler and causing drift in the analytical signal. The sensing fiber was placed in a small volume, stop-flow, stainless steel hybridization chamber (1.5 mm i.d. * 48 mm) which provided a solution volume of 79 µl immediately surrounding the sensing fiber. The temperature of the hybridization cell was controlled by placing the cell in a thermostated housing. The temperature of the solutions in the hybridization cell were accurately determined (±0.2°C) by use of a glass encapsulated thermistor incorporated into the hybridization cell. Solutions containing hybridization buffer, ethidium bromide, and complementary nucleic acid sequences were delivered to the hybridization cell and sensing fiber by use of a peristaltic pump. In all cases, a hybridization buffer/dye solution of 10 mM Tris, 50 mM $MgCl_2$, 2.5×10^{-8} M ethidium bromide at pH 7.3 was used unless otherwise specified. Fluorescence emission from ethidium bromide that was intercalated into immobilized nucleic acid complexes was totally internally reflected within the sensing fiber and directed towards a photomultiplier tube, where the fluorescence intensity could be quantitatively measured. Drift caused by variations in the efficiency of optical coupling, laser intensity and photomultiplier gain were obviated by normalization of all signals to that of a standard

solution of ethidium bromide at 25°C prior to and at the completion of each analysis.

PAGE mobility retardation assay

The solutions of oligonucleotides (10 µl of 30% sucrose/50 mM $MgCl_2$) were incubated at 4°C (96 h) then loaded onto a non-denaturating 15% polyacrylamide gel (90 mM Tris-borate/50 mM $MgCl_2$, pH 8.0). The gels were run at 12.5 mA for 12 h after which the bands in the gel were visualized and photographed by UV illumination followed by ethidium bromide staining.

RESULTS AND DISCUSSION

A goal of this research endeavor was to create a rapid and reliable assay for the detection of triple-helical nucleic acid formation as an extension of work initiated for the detection of duplex formation (14). As a starting point, we chose to investigate the parallel and antiparallel T/AT triplexes as these have been well documented in the literature. Branched nucleic acids as described by Damha *et al.* (3,40) were also used in this study as their unique architecture has been shown to stabilize reversed-Hoogsteen T·AT (antiparallel) triplexes (44). The advantage provided by our optical sensor technology over standard fluorometric work include the low detection limits, reusability and reliability of the device, the non-destructive nature of the assay (where samples may be collected and re-used) and this approach readily lends itself to automation, thereby negating the requirement of highly skilled technicians to carry out the assay.

Figure 4. Fluorescent intensity as a function of temperature dA_{10} functionalized sensors challenged with dT_{10}. Response of the optical sensor to 2.5×10^{-8} M ethidium bromide (solid star).Response of the optical sensor with 5' -> 3'-fiber immobilized dA_{10} to 40 pmol of linear dT_{10} in the presence of 2.5×10^{-8} M ethidium bromide (closed circle). Response of the optical sensor with 3' -> 5'-fiber immobilized dA_{10} to 40 pmol of linear dT_{10} in the presence of 2.5×10^{-8} M ethidium bromide (cross in open circle). Cooling profile of the same nucleic acid system in bulk solution by measurement of absorbance at 260 nm (thick broken line).

Immobilization of oligonucleotides onto optical fibers

The hydroxylated surfaces of the fused silica optical fibers were activated by reaction with GOPS followed by extension with a DMT-HEG linker (Fig. 2). This provides a derivatized surface consisting of a hydrophilic, long-chain spacer arm with a DMT- protected hydroxyl terminus onto which oligonucleotides may be assembled via solid-phase phosphoramidite synthesis (Materials and Methods). This linker was chosen because it is stable to standard oligonucleotide deprotection conditions (37), and provides a fluid environment which facilitates hybridization between immobilized DNA strands and the target strands in solution (47).

Parallel and anti-parallel T-AT triplex considerations

Formation of the intermolecular triplex $2*dT_{10}{:}dA_{10}$ may be characterized by a variety of techniques including UV melting studies, molecular modeling, circular dichroism and NMR spectroscopy (48,49).

In the pyrimidine motif, the third dT_{10} strand interacts by means of Hoogsteen hydrogen bonds with the dA_{10} strand in target duplex, and is oriented parallel to it. In melting experiments (Mg^{2+} buffer), the triplex $2*dT_{10}:dA_{10}$ has two resolved transitions, one for dissociation of the third strand from the duplex, i.e., $dT_{10}*dA_{10}:dT_{10}$ -> dT_{10} + $dA_{10}:dT_{10}$ (T_m = 18°C), and one for dissociation of the duplex into its component strands, i.e., $dA_{10}:dT_{10}$ -> dA_{10} + dT_{10} (T_m = 32°C) (50). Thus association of the third (dT_{10}) strand with the duplex ($dA_{10}:dT_{10}$) is thermodynamically weaker than duplex formation itself (51).

Work done in our laboratories has shown that branched oligonucleotides are useful probes for stabilizing triplex DNA (44). The branched oligomer **1** (Fig. 3) for instance, binds to dA_{10} via reversed-Hoogsteen interactions to give a three-stranded complex in which both dT_{10} strands are antiparallel to the purine (dA_{10}) strand. The formation of this triplex was induced by linkage of two dT_{10} strands through their 5'-ends via coupling to riboadenosine at the neighboring 2' and 3' oxygen atoms. Although this motif had been observed for T-AT bases in complexes dominated by pur·pur:py bonding (e.g., G·GC, A·AT) (52,53), it has only been observed recently for $dT_n:dA_n$ complexes (44,54). Thermal denaturation and renaturation profiles of a mixture of **1** and dA_{10} (1:1) in Mg^{2+} buffer show a single transition from bound to unbound complex (44), consistent with its formation involving a single bimolecular step, i.e., **1** + dA_{10} -> triplex **1**:dA_{10} (T_m = 35°C).

Figure 5. Fluorescent intensity as a function of temperature for the mixed base sequence icosanucleotide functionalized fibers. Upper curve: response of the optical sensor to 20 pmol of linear complement icosanucleotide in the presence of 2.5×10^{-8} M ethidium bromide. Lower curve: response of the optical sensor to 2.5×10^{-8} M ethidium bromide.

Figure 6. Fluorescent intensity as a function of temperature for **1** using reversed orientation 3'-dA_{10}-5'-fiber derivatized sensors. Response of the optical sensor to 40 pmol of **1** in the presence of 2.5×10^{-8} M ethidium bromide (closed circle) and to the 2.5×10^{-8} M ethidium bromide solution alone (solid star). Cooling profile of the same nucleic acid system in bulk solution by measurement of absorbance at 260 nm (broken line).

Triplex studies using derivatized optical fibers with normal (5'-dA_{10}-3'-fiber) oligonucleotide orientation

Decadeoxyadenylic acid (dA_{10}) was grown in the conventional 3' to 5' direction from the fiber surface. Solutions of hybridization buffer containing ethidium bromide, ethidium bromide with dT_{10} or ethidium bromide with **1** were heated (~60°C) in the hybridization chamber containing the decaadenylic acid functionalized optical fibers and renaturation was followed spectroscopically. Fluorescence intensity as a function of temperature for 5'-dA_{10}-3'-fiber functionalized sensors challenged with dT_{10}/ethidium bromide is shown in Figure 4. As the temperature was lowered to 20°C, there was an increase in the fluorescence intensity due to the quantum yield enhancement of the duplex intercalated ethidium bromide.

Upon further cooling, a decrease in the fluorescence intensity with decreasing temperature was observed, indicative of ligand exclusion due to triplex formation ($2*dT_{10}:dA_{10}$). In order to verify that triplex formation was alone responsible for the exclusion of the ethidium cation and the resulting decrease in fluorescence intensity, a control experiment was done using optical fibers functionalized with a20 nt sequence of mixed base composition. Because this sequence lacked a pyrimidine $(Py)_n$ or purine $(Pu)_n$ stretch, only a double-stranded complex could form on the surface of the optical sensor upon binding to a complementary sequence. The hybridization experiment was carried out under the same conditions as for the dA_{10} functionalized fibers with the exception of the hybridization buffer (1 M NaCl, 50 mM PO_4^{2-}, pH 7.0). Intense fluorescence with a negative temperature coefficient was observed for the duplex system over the temperature range studied (10-65°C, T_m = 73°C). The control experiment with ethidium bromide and no complementary oligonucleotide showed a negative temperature coefficient over the same temperature range with no such dramatic increase in intensity (Fig. 5).

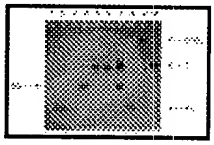


Figure 7. Photograph of a UV-shadowed native polyacrylamide gel containing single strands, duplex and triple helical complexes of branched and linear controls. DNA samples were loaded in 50 mM $MgCl_2$, and 30% sucrose. Lane 4, dT_{10}; lane 5, $dT_{10}:dA_{10}$ (1:1); lane 6, $dT_{10}:dA_{10}$ (2.5:1); lane 7, $dT_{10}:dA_{10}$ (4:1); lane 8, dA_{10}; lane 9, **1** + dA_{10}; lane 10, **1**. As can be noted, the $dT_{10}:dA_{10}$ triplex (lane 7) showed a slight retardation in the mobility relative to the corresponding duplex (lanes 5 and 6). The slowest mobility was observed in lane 9 for the branched triplex **1**:dA_{10}.

Interestingly, upon exposure of the optical sensor to the reversed-Hoogsteen forming **1**, no significant increase in fluorescence intensity over that of the ethidium bromide alone in solution was observed (data not shown). The geometrical constraints of compound **1** are such that, if a complex formed with the immobilized dA_{10} strand in this particular (fiber-3'-dA_{10}-5') orientation, the branch-point riboadenosine moiety would be oriented toward the fiber surface, and thus present a steric barrier to triplex formation. In order to facilitate the formation of the desired antiparallel branched triplex (and test whether steric interference surrounding the branch-point prevented triple-helical formation), an optical sensor having dA_{10} strands in the opposite orientation from the surface, i.e., fiber-5'-dA_{10}-3', was prepared.

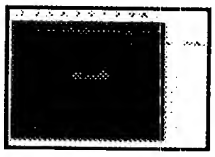


Figure 8. Photograph of an ethidium bromide-stained native polyacrylamide gel (same gel as Fig. 7) containing single strands, duplex and triple helical complexes of branched and linear controls. DNA samples were loaded in 50 mM $MgCl_2$, and 30% sucrose. Lanes 4-10 are the same as those indicated in Figure 7. As can be noted, the $dT_{10}:dA_{10}$ triplex (lane 7) showed a slight retardation in the mobility relative to the corresponding duplex (lanes 5 and 6). The slowest mobility was observed in lane 9 for the branched triplex **1**:dA_{10}. Note that only the duplexes and triplexes showed ethidium bromide fluorescence.

Triplex studies using derivatized optical fibers with reversed (3'-dA_{10}-5'-fiber) oligonucleotide orientation

The fluorescence intensity versus temperature profile with dT_{10} shows an initial increase in fluorescence

intensity with decreasing temperature, indicative of duplex formation (Fig. 4). With further cooling of the system, the polarity of the fluorescence intensity temperature coefficient then inverts, indicative of triplex formation. Treatment of the optical sensor with **1** also provided a fluorescence intensity with a positive temperature coefficient at temperatures below the T_m (35°C), indicative of the formation of the reverse-Hoogsteen complex (Fig. 8).

The results of these experiments can be best understood by considering the two key competing factors which influence the net fluorescence intensity temperature coefficient. Firstly, the fluorescence quantum efficiency of the intercalant ligand bound to triple-stranded nucleic acids is greater than that of the ligand bound to double-stranded nucleic acid (36,45,46). This is the result of the triple-stranded structure being more rigid than the double-stranded nucleic acid structure, thereby providing superior shielding of the intercalated fluorophore from non-radiative collisional deactivation. In both cases, triplex and duplex, the quantum efficiency of the bound fluorophore increases with decreasing temperature (i.e., displays a negative temperature coefficient) owing to the overall reduction in the molecular motion in the system. The second factor influencing the net fluorescence emission is the binding efficiency of the intercalant ligand to each substrate type. Not as many ethidium cations can be accommodated per base triplet as per base pair. In addition, further exclusion of ethidium cation occurs with decreasing temperature in triple-helical nucleic acids, thereby providing a fluorescence intensity with a positive temperature coefficient. At low temperatures, the exclusion process dominates the fluorescence signal, thereby providing a means for elucidation of triple-strand formation.

In greater detail, it can be inferred from the data of Scaria and Shafer (36) that under these conditions of ionic strength and pH, a temperature below 25°C is required for the ethidium cation exclusion process to dominate the net fluorescence signal. Given that intercalation occurs at a maximum of every 2.8 base triplets and once per 2.4 base pairs at 25°C, a 14% reduction in the amount of intercalated ethidium occurs upon triple-strand formation. However, within the triplex structure, the fluorescence quantum yield of the remaining intercalated ethidium cation increases by 19% for the S_1 -> S_0 electronic transition, thereby resulting in a net fluorescence intensity change of +2.3%. Therefore, direct correlation between the T_m for triplex formation and the onset of fluorescence emission with a positive temperature coefficient will be observed for systems of nucleic acids which have T_m values at or below ~25°C. This is consistent with our findings (Fig. 4) whereby the decrease in fluorescence intensity from the sensor correlates well with the temperature at which dT_{10} associates to the dT_{10}/dA_{10} duplex (T_m = 18°C). Although the transition for triple-strand formation between **1** and the immobilized dA_{10} occurs at 35°C (Fig. 6), a decrease in fluorescence intensity was not observed until the system was cooled to below ~25°C. In this regard, our fluorescence studies involving ethidium bromide binding to triple-helices is in full agreement with several earlier findings. Our system is then limited in terms of being able to identify the duplex to triplex transition temperature for nucleic acid systems with T_m values at or below 25°C. This does not, however, limit the applicability of this technology in terms of being a useful strategy to identify triplex formation.

It is also interesting to note in Figure 6, for the binding of **1** with immobilized dA_{10}, a significant fluorescence intensity is observed over the temperature range from ~50 to 60°C. This is indicative of the presence of intercalated ethidium cation. This is contrary to data presented in the UV denaturation/renaturation profiles for the same oligonucleotide system in solution where no significant quantity of complex formation exists over that temperature regime. A possible explanation for this unusual observation is that the ionic strength at or near the surface of the optical sensor may be greater

than that of bulk solution owing to the presence of the immobilized polyanionic nucleic acid strands. As such, a shift in the T_m to higher temperatures would be expected. This is consistent with our previously reported data where binding of dA_{20} to immobilized dT_{20} was found to have a T_m value greater than that of the same oligonucleotide system in solution (14).

PAGE mobility retardation assay

Gel-shift experiments confirmed the interaction of ethidium bromide with the complexes observed in these studies. The electrophoretic mobility of the dT_{10}:dA_{10} duplex, both the Hoogsteen and reverse-Hoogsteen paired T·AT triplexes, and that of their component strands, was studied at 4°C. Following electrophoresis, the gels were visualized by UV shadowing, and by staining with ethidium bromide (Figs 7 and 8, respectively). The Hoogsteen triplex migrated more slowly than the duplex while the reversed-Hoogsteen triplex showed the slowest mobility of all, which is characteristic of branched nucleic acid structures (55). Association of **1** and dA_{10} was quantitative as evidenced by the complete disappearance of compound **1** and dA_{10}, when mixed in equimolar amounts, as visualized in the gel (Fig. 7). The stoichiometry of interaction between dT_{10} and dA_{10} for the duplex and Hoogsteen triplex was also confirmed by studies at different concentrations of the two oligonucleotides. When stained with ethidium bromide and illuminated by a UV lamp, fluorescence was observed only in the bands corresponding to the complexes, not single strands (Fig. 8). This is consistent with the well-known intercalative binding motif of ethidium bromide (56). As previously suggested by the biosensor studies, the **1**/dA_{10} reverse-Hoogsteen triplex gave the lowest fluorescence intensity, which could be caused by the limited availability of ethidium binding sites in this complex.

Conclusions

In conclusion, a novel method for the detection of triple-helical nucleic acid formation has been demonstrated. The complementary oligonucleotide, dT_{10}, which was expected to hybridize via a double-stranded Watson-Crick motif to immobilized dA_{10} provided a fluorescence intensity with a negative temperature coefficient upon cooling the system below the duplex melting temperature (T_m = 32°C). Upon further cooling, to the point where Hoogsteen T*AT triple-helix formation occurred, a fluorescence intensity change with a positive temperature coefficient was observed as a result of exclusion of the ligand from the triplex structure. Similar results were observed for triplex formation between dT_{10} and the immobilized dA_{10} sequence in both the normal (fiber-3' -> 5') orientation and the reversed (fiber-5' -> 3') orientation. The reversed-Hoogsteen T·AT triplex formed with **1** and the immobilized dA_{10} grown in reversed orientation (fibre-5' -> 3') also provided a fluorescence intensity with a positive temperature coefficient, consistent with triplex formation and ligand exclusion. Correlation between the triplex T_m and the temperature at which the temperature coefficient of the fluorescence intensity changes from negative to positive may be observed for nucleic acid systems with a triplex T_m below ~25°C. Determination of triplex formation may be done rapidly (in minutes) by setting the initial temperature of the system to that of the triplex T_m and then slowly cooling the system (-0.5°C/min) for a few minutes to determine the fluorescence intensity temperature coefficient.

Further studies will be directed to expanding the triple-helix sequence context, investigations of mismatch sensitivity, and developing less limiting fluorescent dyes. Optical sensors with covalently bound intercalant have been created in our laboratories which provide a reagentless sensing system with fast

response times (<6 min for full analytical response) for double-strand formation. Investigations of triplex formation on these reagentless sensors will also be evaluated in diagnostic assays, as they eliminate the problem of doubled-stranded DNA in the sample solution (e.g., in a biological sample) procuring all of the intercalant present in the buffer solution.

ACKNOWLEDGEMENTS

We acknowledge the Natural Science and Engineering Council of Canada (NSERC), and Le Fonds pour la Formation des Chercheurs et l'Aide à la Recherche (FCAR) for support of this work. We also acknowledge additional support from NSERC for a postgraduate scholarship awarded to P.A.E.P. and a postdoctoral scholarship awarded to R.H.E.H. We would also like to acknowledge Sebastien Robidoux for supplying the thermal renaturation data presented herein.

REFERENCES

1 Thomas, D. (1995) *BioTechnology*, **13**, 474-481.

2 Laval, J.M., Chopineau, J. and Thomas, D. (1995) *Trends Biotechnol.*, **13**, 474-481.

3 Hudson, R.H.E. and Damha, M.J. (1993) *J. Am. Chem. Soc.*, **115**, 2119-2124.

4 Horn, T. and Urdea, M.S. (1989) *Nucleic Acids Res.*, **17**, 6959-6967. MEDLINE Abstract

5 Kool, E.T. (1996) *Annu. Rev. Biophys. Biomolec. Struct.*, **25**, 1-28.

6 Chen, G. and Seeman, N.C. (1991) *Nature*, **350**, 631-633.

7 Ray, R.A., Cooper, P.J. and Wakefield, A.J. (1995) *BioTechnology*, **13**, 445-447.

8 McGown, L.B., Joseph, M.J., Pitner, J.B., Vonk, G.P. and Linn, C.P. (1995) *Anal. Chem.*, **67**, A663-A668.

9 Millan, K.M., Saraullo, A. and Mikkelsen, S.R. (1994) *Anal. Chem.*, **66**, 2943-2948. MEDLINE Abstract

10 Su, H.B., Chong, S. and Thompson, M. (1996) *Langmuir*, **12**, 2247-2255.

11 Nilsson, P., Persson, B., Uhlén, M. and Nygren, P. (1995) *Anal. Biochem.*,**22**, 400-408.

12 Bates, P.J., Dosanjih, H.S., Kumar, S. Jenkins, T.C., Laughton, C.A. and Neidle, S. (1995) *Nucleic Acids Res.*, **23**, 3627-3632. MEDLINE Abstract

13 Mirkin, C.A., Letsinger, R.L., Mucic, R.C. and Storhoff, J. (1996) *Nature*,**382**, 607-609. MEDLINE Abstract

14 Piunno, P.A.E., Krull, U.J., Hudson, R.H.E., Damha, M.J. and Cohen, H. (1995) *Anal. Chem.*, **67**, 2635-2643.

15 Graham, C.R., Leslie, D. and Squirrell, D.J. (1992) *Biosens. Bioelectron.*, **7**, 487-493. MEDLINE Abstract

16 Abel, A.P., Weller, M.G., Duveneck, G.L., Ehrat, M. and Widmer, H.M. (1996) *Anal. Chem.*, **68**,

2905-2912. MEDLINE Abstract

17 Strobel, S.A., Doucettestamm, L.A., Riba, L., Housman, D.E. and Dervan, P.B. (1991) *Science*, **254**, 1639-1642. MEDLINE Abstract

18 Maher, L.J., Dervan, P.B. and Wold, B. (1992) *Biochemistry*, **31,** 70-81. MEDLINE Abstract

19 Duvalvalentin, G., Thoung, N.T. and Hélène, C. (1992) *Proc. Nat. Acad. Sci. USA*,**89**, 504-508.

20 Stull, R.A. and Szoka, F.C.,Jr (1995) *Pharmceut. Res.*, **12**, 465-483.

21 Ito, T., Smith, C.L. and Cantor, C.R.(1992) *Proc. Natl. Acad. Sci. USA*, **89**, 495-498. MEDLINE Abstract

22 Wang, S.H., Friedman, A.E., Kool, E.T. (1995) *Biochemistry* **34**, 9774-9784.

23 Hoogsteen, K. (1959) *Acta Crystallogr.*, **12**, 822-823.

24 Felsenfeld, G. Davies, D.R. and Rich, A. (1957) *J. Am. Chem. Soc.*, **79**, 2023-2024.

25 Moser, H.E. and Dervan, P.B. (1987) *Science*, **238**, 645-650. MEDLINE Abstract

26 Praseuth, D., Perrouault, L., Doan, T.L., Chassignol, M., Thuong, N. and Hélène, C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1349-1353. MEDLINE Abstract

27 Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J. and Hogan, M.E. (1988) *Science*, **241**, 456-459. MEDLINE Abstract

28 Beal, P.A. and Dervan, P.B. (1991) *Science*, **251**, 1360-1363. MEDLINE Abstract

29 Miller, J.H. and Sobell, H.M. (1966) *Proc. Natl. Acad. Sci. USA*, **55**, 1201-1205. MEDLINE Abstract

30 Morgan, A.R. and Wells, R.D. (1968) *J. Mol. Biol.*, **37**, 63-80. MEDLINE Abstract

31 Lee, J.S., Johnson, D.A. and Moran, A.R. (1979) *Nucleic Acids Res.*, **6**, 3073-3091. MEDLINE Abstract

32 Hoogsteen, K. (1963) *Acta Crystallogr.* **16**, 907-916.

33 Waring, M.J. (1966)*Biochem. J.*, **143**, 483-486.

34 LePecq, J.B. and Paoletti, C. (1967) *J. Mol. Biol.*, **27**, 87-106. MEDLINE Abstract

35 Mergny, J.L., Collier, D., Rougée, M., Montenay-Garestier, T. and Hélène, C. (1991) *Nucleic Acids Res.*, **19**, 1521-1526. MEDLINE Abstract

36 Scaria, P.V. and Shafer, R.H. (1991) *J. Biol. Chem.*, **266**, 5417-5423. MEDLINE Abstract

37 Maskos, U. and Southern, E.M. (1992) *Nucleic Acids Res.*, **20**, 1679-1684. MEDLINE Abstract

38 Pon, R.T. (1993) In Agrawal, S. (ed.) *Methods in Molecular Biology-Protocols for Oligonucleotides and Analogs*. Humana Press, Totowa, NJ, Vol 20, 465-496.

39 Van de Sande, J.H., Ramsing, N.B., Germann, M.W., Elhorst, W., Kalisch, B.W., Kitzing, E.V., Pon,

R.T., Clegg, R.C. and Jovin, T.M. (1988) *Science* **241**, 551-557. MEDLINE Abstract

40 Damha, M.J., Ganeshan, K., Hudson, R.H.E. and Zabarylo, S.V. (1992) *Nucleic Acids Res.*, **20**, 6565-6573. MEDLINE Abstract

41 Damha, M.J. and Ogilvie, K.K., (1993) Agrawal, S. (ed.) *Methods in Molecular Biology-Protocols for Oligonucleotides and Analogs*. Humana Press, Totowa, NJ, Vol 20, pp 81-114.

42 Kibler-Herzog, L., Zon, G., Whittierm, G., Shaikh, M. and Wilson, W.D. (1993) *Anti-Cancer Drug Des.*, **8**, 65-79.

43 Brennan, J.D., Brown, R.S., McClintock, C.P. and Krull, U.J. (1991) *Anal. Chim. Acta*,**237**, 253-263.

44 Hudson, R.H.E., Uddin, A. and Damha, M.J. (1995) *J. Am. Chem. Soc.*, **117**, 12470-12477.

45 Salunkhe, M., Wu, T. and Letsinger, R.L. (1992) *J. Am. Chem. Soc.*, **114**, 8768-8772.

46 Fang, Y., Bai, C.L., Zhang, P.C., Cao, E.H. and Tang, Y.Q. (1994) *Science In China (Ser. B)*,**37**, 1306-1312.

47 Wolf, S.F., Haines, L., Fisch, J., Kremsky, J.N., Dougherty, J.P. and Jacobs, K. (1987) *Nucleic Acids Res.*,**15**, 2911-2926. MEDLINE Abstract

48 Hopkins, H.P., Hamilton, D.D.Wilson, W.D. and Zon, G. (1993) *J. Phys. Chem.*, **24**, 6555-6563.

49 Radhakrishnan, I., Patel, D.J., Priestly, E.S., Nash, H.M. and Dervan P.B. (1993) *Biochemistry*, **32**, 11228-11234. MEDLINE Abstract

50 Pilch, D., Levenson, C., Shafer, R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1942-1946. MEDLINE Abstract

51 Pilch, D.S., Brousseau, R. and Shafer, R.H. (1990) *Nucleic Acids Res.*, **18**, 5743-5750. MEDLINE Abstract

52 Fox, K.R. (1994) *Nucleic Acids Res.*, **22**, 2016-2021. MEDLINE Abstract

53 Cassidy, S.A., Strekowski, L., Wilson, W.D., and Fox, K.R. (1994) *Biochemistry*, **33**, 15338-15347. MEDLINE Abstract

54 Dagneaux, C., Liquier, J. and Taillandier, E. (1995) *Biochemistry*, **34**, 14815-14818. MEDLINE Abstract

55 Wallace, J.C. and Edmons, M. (1983) *Proc. Natl. Acad. Sci. USA*,**80**, 950-954. MEDLINE Abstract

56 Lim, C.S. (1994) *BioTechniques*, **17**, 626-628. MEDLINE Abstract

---

*To whom correspondence should be addressed. (M.J.D.) Tel: +1 514 398 7552; Fax: +1 514 398 3797; Email: damha@omc.lan.mcgill.ca(U.J.K.) Tel: +1 905 828 5437; Fax: +1 905 828 5425; E-mail: ukrull@credit.erin.utoronto.ca

+Present address: Department of Chemistry, University of Western Ontario, London, Ontario, Canada

The authors wish it to be known that, in their opinion, the first two authors should be considered as joint

First Authors